

Journalistische Aktualität im Internet. Ein Experiment mit den „News-Suchfunktionen“ von Suchmaschinen

Der Artikel thematisiert die Suche nach aktuellen, journalistischen Inhalten im Internet mithilfe von Suchmaschinen und deren News-Suchfunktionen. Suchmaschinen nehmen im undurchsichtigen Internet-Nachrichtendschungel Gatekeeper-Rollen ein und tragen demnach dieselbe Verantwortung wie Journalisten in den klassischen Medien. Obwohl Suchmaschinen einen hohen Standard erreicht haben, wird die Qualität der Suchergebnisse durch inhaltliche, ökonomische und technologische Aspekte eingeschränkt. Suchmaschinen bleiben für die Wissenschaft Black Boxes. Sie legen weder ihre Quellen noch ihre Ranking-Methoden offen und machen zudem selten Angaben über Veröffentlichungsdaten der Nachrichten. Drei Experimente, die in verschiedenen Zeitintervallen durchgeführt wurden, haben die Suche nach gültigen aktuellen Suchergebnissen analysiert. Nur drei Suchmaschinen nutzen für ihre Suche eine Reihe von unabhängigen Quellen: Alta Vista, Google und Paperball. Auch diese garantieren jedoch keine vollständige Wiedergabe der Nachrichtenlage.

1 Einleitung

Die Entwicklung der Erschließung von journalistischen Nachrichteninhalten in Suchmaschinen lässt sich in vier Phasen einteilen:

In der ersten Phase bestand keine Notwendigkeit, Nachrichteninhalte in Suchmaschinen einzubinden. Journalistische Online-Nachrichtenangebote waren noch nicht in großer Zahl vorhanden. Die Anstrengungen der Suchmaschinen konzentrierten sich auf die Erfassung konventioneller Webseiten, die keiner hohen Aktualisierungsfrequenz unterlagen.

Die zweite Phase ist durch die Umwandlung der damals großen Suchmaschinen in Portale¹ gekennzeichnet, wie beispielsweise im Fall von AltaVista

1 „(...) Bezeichnung für allgemeine Einstiegspunkte für Nutzer des WWW, die nach bestimmten Inhalten suchen und dafür auf vorsortierte Suchkataloge zurückgreifen“ (Klußmann 2001: 777).

oder im Bereich der Web-Verzeichnisse bei Yahoo². Da Nachrichteninhalte als integraler Bestandteil der allgemeinen Portale angesehen wurden (Rösch 2001b: 28), waren sie in der Regel durch einen Zukauf von Agenturen eingebunden.

Die dritte Phase kennzeichnet sich durch die Rückkehr der Suchmaschinen zu ihrer Kernkompetenz, der Suche selbst. Nach dem Vorbild von Google boten Suchmaschinen auf ihrer Homepage in erster Linie einfache Suchformulare an. Die Suche selbst fragte einzig den regulären Datenbestand der Suchmaschine ab. Durch Aktualisierungszyklen von etwa vier Wochen (vgl. Notess 2003) konnten aktuelle Nachrichten allerdings nicht zeitnah erfasst werden. Zwar kreuzten die Crawler³ der Suchmaschinen permanent durch das WWW; die gefundenen Seiten konnten jedoch nicht in Echtzeit in den Index integriert werden, sondern wurden erst bei einem vollständigen Update des Index suchbar. Gleichzeitig stieg das Angebot an aktuellen Nachrichten im Internet stark an.⁴

Mittlerweile ist die Nachrichtensuche in der vierten Phase angelangt. Das große Nachrichtenangebot wird nun von den Suchmaschinen erfasst, alle großen Suchmaschinen verfügen inzwischen über eigene Funktionen zur Nachrichtensuche. Des Weiteren wird die vierte Phase durch die zunehmende Hinwendung der Nutzer zu Nachrichteninhalten gekennzeichnet. Angebote mit tagesaktuellen Informationen werden von deutschen Internet-Nutzern derzeit am häufigsten aufgerufen. Außerdem steigt das Interesse an News-Lettern und News-Tickern. Beide Angebotsformen ersetzen für einige Nutzer inzwischen den Kauf einer Tageszeitung (van Eimeren/Gerhard/Frees 2003: 348).

1.1 Gliederung und Fragestellung

Im deutschsprachigen Raum sind News-Suchmaschinen und ihre Funktionalität bisher kaum erforscht worden. Die in diesem Band mehrfach angesprochenen Gatekeeper-Effekte von Suchmaschinen wirken sich jedoch insbesondere bei den News-Suchmaschinen aus; die Strukturierungsmacht ist hier angesichts der aktuellen journalistischen Informationen besonders sensibel zu betrachten. Es ergaben sich daher zahlreiche Fragen, die im Rahmen einer explorativen Pilotstudie zu stellen sind. Folgende Aspekte stehen im Mittelpunkt:

2 Die URLs aller behandelten Suchmaschinen sind im Anhang angegeben.

3 Ein Crawler ist ein Softwareprogramm, das automatisch Seiten des WWW aufruft, ihren Inhalt entsprechend bestimmter Algorithmen erfasst und die Ergebnisse in die Datenbank („Index“) einer Suchmaschine einfügt (vgl. Brin/Page 1998).

4 Dass die Suchmaschinen nicht mehr schnell genug auf aktuelle Nachrichtenangebote reagieren konnten, zeigte sich besonders deutlich im Umfeld der Ereignisse des 11. September 2001. Wiggins (2001) beschreibt den Umgang der Suchmaschine Google mit dem erhöhten Bedarf an Nachrichteninhalten in dieser Zeit: Erst verwies Google auf seiner Homepage auf Radio und Fernsehen als Informationsquellen, dann wurde eine manuelle Linkliste mit Verweisen auf wichtige Nachrichten-Websites angeboten.

- Werden aktuelle journalistische Inhalte von Suchmaschinen zeitnah erfasst und angezeigt? Wie lange dauert es, bis ein neuer Inhalt angezeigt wird?
- Wird in der Suchmaschine und/oder auf der Seite des Anbieters eindeutig gekennzeichnet, wann eine Nachricht geschrieben wurde? Wird immer die aktuelle Version eines Beitrags angezeigt?
- Welche speziellen Suchfunktionen für journalistische Inhalte gibt es innerhalb und außerhalb der bestehenden Suchmaschinen (z.B. Google-News, Paperazzi)? Wie sind diese in das bestehende Angebot integriert? Werden Nachrichten von anderen Angeboten – insbesondere von PR-Texten – getrennt?
- Gibt es nachvollziehbare Kriterien für das Ranking der einzelnen Beiträge, zum Beispiel hinsichtlich Aktualität, Relevanz oder Verbreitungsgebiet? Welche Anbieter werden erfasst? Werden bestimmte Titel immer wieder bevorzugt, zum Beispiel weil sie Kooperationspartner des Suchmaschinenbetreibers sind?
- Gibt es (technische) Möglichkeiten, mit denen die Anbieter von journalistischen Inhalten sicherstellen können, dass ihre aktuelle Berichterstattung von den Suchmaschinen auch rechtzeitig gefunden und angezeigt wird?

Nach einem kurzen Überblick über News-Suchmaschinen im Internet beschreibt der Artikel die Technologien und Gesichtspunkte, die die Suche nach nachrichtlichen Inhalten hinlänglich beeinflussen. Quellentransparenz und die Beziehung zwischen Journalismus, Werbung und PR sind dabei zentrale Aspekte, bevor verschiedene deutschsprachige Suchmaschinen vorgestellt werden. An die Diskussion über die spezifischen Herausforderungen von Nachrichtensuchen, die auf Aktualität und Ranking basieren, schließen sich drei Experimente an, die die jeweilige Leistung der News-Suchmaschinen evaluieren: Der Fünf-Tage-Test analysiert, in welchem Umfang Suchmaschinen relevante und aktuelle Ergebnisse über ein Ereignis liefern. Der 24-Stunden-Test gibt Aufschlüsse über den Zeitpunkt der Veröffentlichung in Suchmaschinen. Der Reaktionstest untersucht die Geschwindigkeit, mit der aktuelle Vorkommnisse in den Nachrichtenindex von Suchmaschinen integriert werden. Eine Diskussion der Ergebnisse und der gegenwärtigen Bedingungen von Nachrichtensuchen im Internet fasst die Studie zusammen.

2 Nachrichtensuchmaschinen im Überblick

2.1 *Technologie und Probleme der Nachrichtensuchmaschinen*

2.1.1 Spezielle News-Suchmaschinen

Die spezielle Funktion der Nachrichtensuche wird von zwei Typen von Suchmaschinen ausgeführt: Eigenständige Nachrichtensuchmaschinen, die nicht an eine Universalsuchmaschine gebunden sind, werden ausschließlich zum Zweck der Nachrichtensuche betrieben (z.B. Paperazzi). Ihnen gegenüber stehen spezielle Suchfunktionen für Nachrichten innerhalb der Websuchmaschinen (z.B. Google-News als Teil von Google). Aus technischer Sicht gibt es jedoch kaum Unterschiede. Die Nachrichtensuche bei Websuchmaschinen ist in der Regel nicht in den normalen Webseitenindex integriert und greift stattdessen auf einen eigenen Datenbestand zurück. Nur auf der Ebene der Benutzeroberfläche sind Websuche und Nachrichtensuche integriert, ähnlich wie auch die Bilder- und Audio-/Video-Suche.

2.1.2 Technologie der Nachrichtensuche

Prinzipiell lassen sich zwei Verfahren bei der Erschließung von Nachrichten durch Suchmaschinen unterscheiden. Entweder verwendet die Suchmaschine einen zugekauften Nachrichtenbestand, der etwa von einer Nachrichtenagentur geliefert wird (so genannte Newsfeeds), oder die Suchmaschine erfasst selbst die Inhalte von Nachrichtenseiten mithilfe eines Crawlers.

Im ersten Fall ist die technische Lösung relativ unkompliziert. Die Nachrichten werden in strukturierter Form geliefert. Diese formale Unterteilung muss pro Zulieferer nur einmal an die Struktur der von der Suchmaschine verwendeten Datenbank angepasst werden. Allerdings werden bei der Erschließung in der Regel nur wenige Quellen abgedeckt. Die Gründe sind vor allem wirtschaftlicher Natur: Für die zugelieferten Nachrichten müssen die Suchmaschinenbetreiber bezahlen.

Erschließt die Suchmaschine die Inhalte mittels Crawling, ergibt sich das Problem der unterschiedlichen Strukturierung von Nachrichten-Websites. Die Suchmaschine muss erkennen, bei welchem Teil der jeweiligen HTML-Seite es sich um journalistisch bzw. nachrichtlich relevante Inhalte handelt. Texte innerhalb der Navigationselemente und in Verweisen auf weitere Artikel sollen nicht in den Datenbestand der Suchmaschine aufgenommen werden; sie würden die Suchergebnisse verfälschen (vgl. auch 2.3.1).

Ein anderes aus der Weberschließung bekanntes Problem entfällt dafür: Die Suchmaschine muss keine unerwünschten Inhalte (Spam)⁵ ausschließen,

⁵ Zum Einsatz von Spam als Mittel zur Manipulation von Suchmaschinen: Machill et al. 2003: 125 ff.; Karzauninkat 2003: 528 ff.

da die Nachrichtenseiten, die in die Erschließung eingehen, bei allen Betreibern von Hand ausgewählt werden. Damit wird eine gewisse Qualitätssicherung gewährleistet. Die Frage der Belastung der Suchmaschine durch ständiges Crawlen ist zu vernachlässigen⁶, da es sich bei den Nachrichtenquellen um eine relativ geringe Anzahl von Sites handelt.

Ein zentrales Element im Bereich der Suchmaschinenteknologie ist das Ranking⁷. Wird für das Ranking nur der Text der gefundenen Meldungen ausgewertet, werden also allein anfrageabhängige Faktoren (Query-dependent Factors) berücksichtigt, so haben alle gefundenen Nachrichten potenziell die gleiche Chance, bei Suchanfragen an erster Stelle der Trefferliste platziert zu werden. Es wird nicht berücksichtigt, dass einige Nachrichtenquellen von den Nutzern als bedeutender eingestuft werden als andere.

Eine Bevorzugung von „Top-Quellen“ wird bei einigen Nachrichtensuchmaschinen mittels eines Rankings durch anfrageunabhängige Faktoren (Query-independent Factors) erreicht. Dabei findet eine automatisierte Bewertung der Qualität der Sites statt, die sich später auf das Ranking der einzelnen Seiten⁸ dieser Angebote auswirkt. Die journalistische Qualität ist nicht unbedingt von Interesse.

Bei den Nutzern mag mitunter der Eindruck entstehen, mit einer speziellen Nachrichtensuchmaschine ließe sich eine vollständige Presseschau erstellen. Dies ist jedoch weiterhin nur unter Zuhilfenahme professioneller Datenbanken möglich. Allerdings können Nachrichtensuchmaschinen unterstützend eingesetzt werden, da sie auch Quellen integrieren, die nicht über kommerzielle Datenbanken vermarktet werden oder nur als Online-Angebote vorhanden sind.

2.1.3 Integration der Nachrichtensuche in Universalsuchmaschinen

Im Idealfall würden Nachrichteninhalte kontinuierlich in den Datenbestand normaler Suchmaschinen eingespeist und damit recherchierbar werden. Diesem Verfahren stehen jedoch technische Hürden entgegen: Die Technik der meisten Suchmaschinen erlaubt es nicht, alle gefundenen Dokumente direkt in die Datenbank der Suchmaschinen zu integrieren. Erst bei einer vollständigen Indexaktualisierung können die neuen Dokumente aufgenommen werden. Da die Datenbestände der Suchmaschinen enorm groß sind, ist eine vollständige Indexaktualisierung sehr aufwendig (vgl. Bekavac 2002: 2.3.1.3.;

6 Bereits der Prototyp von Google war 1998 mit vier Crawlern in der Lage, 100 Webseiten pro Sekunde zu durchsuchen (Brin/Page 1998).

7 Die Ergebnisse einer Suchanfrage müssen in eine Rangfolge gebracht werden. Diese richtet sich nach unterschiedlichen Kriterien, z.B. der (von der Suchmaschine mittels unterschiedlicher Methoden) angenommenen Relevanz der Seite für die Suchanfrage, der Aktualität oder aber nach dem Grad der Verlinkung (z.B. „PageRank“).

8 Mit „Site“ wird die Gesamtheit einer Webseite bezeichnet, während unter dem Begriff „Seite“ eine bestimmte HTML-Seite verstanden wird, die Teil einer Site ist (Klußmann 2001: 891).

Wimmeroth/Brochhagen 2003: 160f.). Die Suchmaschine Fireball nimmt Nachrichteninhalte hingegen aufgrund der Möglichkeit der permanenten Indexaktualisierung ohne vollständiges Update sofort in den regulären Webindex auf.

Die zweite technische Hürde sind die von den Suchmaschinen eingesetzten linktopologischen Verfahren wie PageRank. Seiten, die eine größere Verlinkungsdichte aufweisen, werden als bedeutender angesehen als solche mit schwacher oder nicht vorhandener Verlinkung (vgl. Page et al. 1998: 3ff.). Demnach würden aktuelle Nachrichten, die noch keine Links auf sich gezogen haben, selbst bei einer möglichen direkten Integration in den Index nicht auf den ersten Plätzen der Trefferlisten angezeigt werden. Um die Auswahl auszubalancieren, trennen die meisten Suchmaschinen die Nachrichtensuche von der Websuche und den weiteren Spezialsuchen wie Multimedia oder FTP⁹.

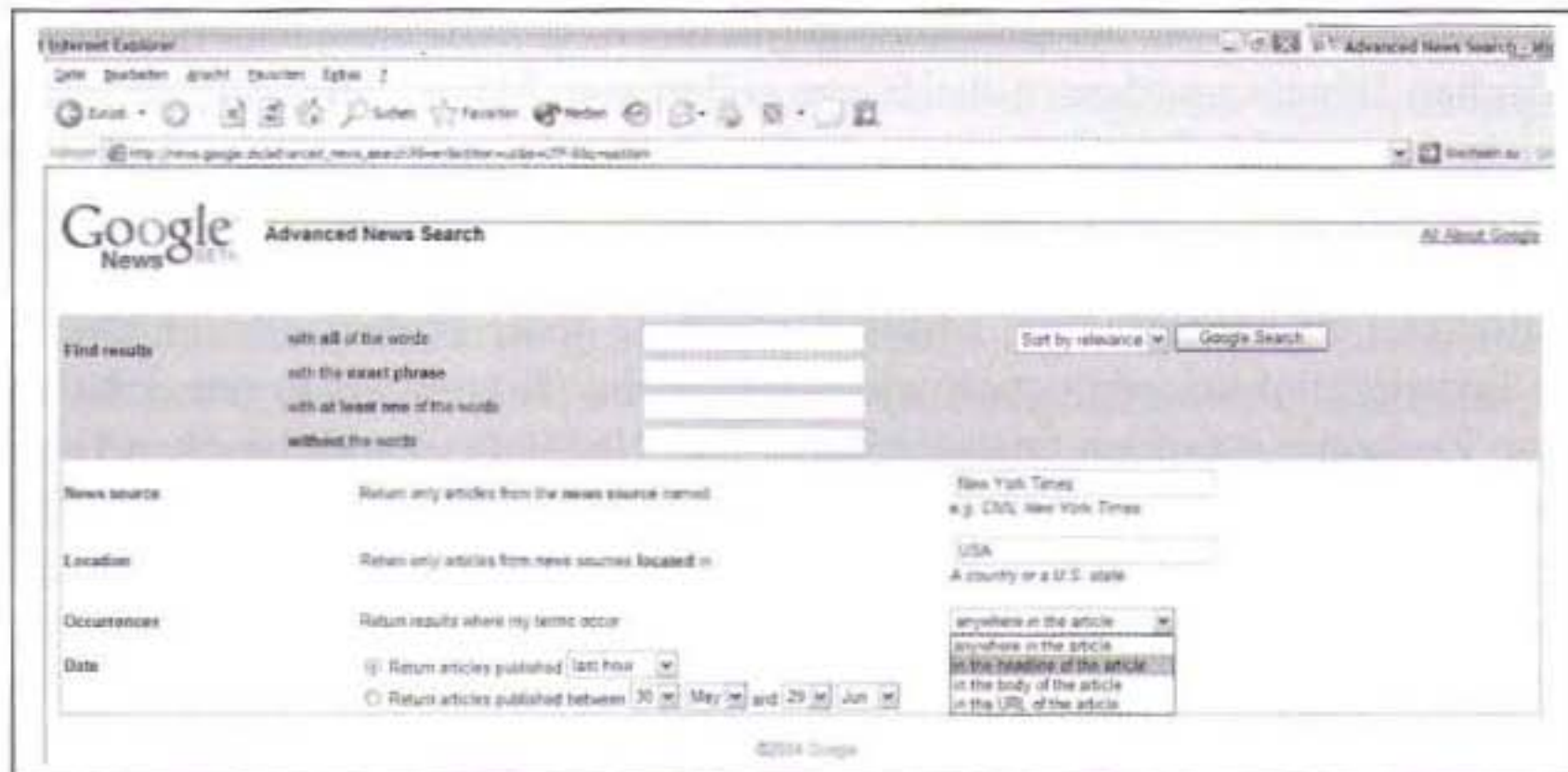
Jeder dieser Spezialsuchen liegt eine eigene Datenbank zugrunde. Zwischen diesen verschiedenen Indizes kann in der Regel durch das Anklicken von so genannten Tabs (Reitern) gewechselt werden. Um diejenigen Nutzer, die nachrichtenrelevante Suchanfragen im Web stellen, trotzdem zu den Treffern zu führen, werden solche meist oberhalb der regulären Ergebnislisten angezeigt. Die vor den Trefferlisten in der Websuche angezeigten Inhalte werden aus einem eigenen Nachrichtenindex generiert. Beim Abschicken einer regulären Suchanfrage werden demnach mindestens zwei Indizes abgefragt. In der Regel kommen noch weitere Datenbanken, etwa die der bezahlten Werbung, hinzu. Bisher werden in den speziellen Nachrichtensuchen der Universalsuchmaschinen allerdings keine kontextbezogenen Werbetreffer angezeigt. Diese erscheinen nur bei Trefferlisten der regulären Websuche.

2.1.4 Lokalisierung

Lokalisierungsfunktionen – ähnlich denen in allgemeinen Suchmaschinen (vgl. Lewandowski 2004 und den Beitrag von Susanne Päch in diesem Band) – werden sich künftig auch in Nachrichtensuchmaschinen entwickeln, teils sind sie bereits vorhanden. So sind bei Paperball die Quellen auch nach Bundesländern erfasst und in der US-Ausgabe von Google-News lässt sich die Suche nach US-Bundesstaaten einschränken (Abb. 1). Denkbar wäre für die Zukunft eine lokale Suche in den relevanten Quellen, die im Umkreis des Nutzers erscheinen bzw. Themen aus diesem Umkreis behandeln.

9 FTP – File Transfer Protocol; spezielle Suche nach „herunterladbaren“ Dateien im Netz.

Abb. 1: Möglichkeiten der lokalisierten Suche unter news.google.com



2.2 Suche nach Nachrichteninhalten

Im Folgenden werden die Suche nach Nachrichteninhalten in den wichtigsten Universalsuchmaschinen sowie die wichtigsten spezialisierten Nachrichtensuchmaschinen für deutschsprachige Inhalte betrachtet: Im Bereich der Universalsuchmaschinen sind dies: GoogleNews, Alltheweb und AltaVista. Ergänzend wurde das Angebot des Portalbetreibers Yahoo aufgenommen, da hier ein guter Zugriff auf Agenturmeldungen besteht, die von den anderen Nachrichtensuchmaschinen nicht angeboten werden. Die deutschsprachigen Nachrichtensuchmaschinen werden durch Paperball repräsentiert, der Nachrichtensuchmaschine von Lycos und Fireball. Als einziges englischsprachiges Angebot wurde MSN Newsbot von Microsoft aufgenommen. Diese Nachrichtensuchmaschine war bis Redaktionsschluss dieses Beitrags (Ende April 2004) noch nicht in einer deutschsprachigen Version verfügbar. Des Weiteren wurden die Internet-Auftritte von Web.de, AOL und T-Online sowie der Metasuchmaschine Metager in die Betrachtung einbezogen. Diese sind nicht als reine Nachrichtensuchmaschinen zu werten, bieten aber ebenfalls entweder Nachrichtenangebote im Zuge ihrer Portalfunktion an bzw. werden im Zuge der eigentlichen Websuche zur Nachrichtensuche verwendet. Zudem besitzen sie aufgrund ihrer Nutzerzahlen eine gewisse Relevanz (siehe auch 2.2.3).

2.2.1 Quellenspektrum und journalistische Professionalität

Das Spektrum der im Internet veröffentlichten Nachrichten ist breit. Das Angebot reicht von den Nachrichtenangeboten der großen Zeitungen, Zeitschriften und Fernsehsender bis hin zu Branchen-Newslettern und von Pri-

vatpersonen betreuten Weblogs (Blogs).¹⁰ Die Nachrichtensuchmaschinen konzentrieren sich bei der Erschließung in der Regel auf die „echten“ journalistischen Inhalte und lassen die Blogs außen vor. Manche Nachrichtensuchmaschinen erschließen jedoch auch Pressemitteilungen von Unternehmen und Verbänden.

Der Journalismus verliert zunehmend sein Monopol, den Zugang zur Öffentlichkeit zu kontrollieren. Mittlerweile kann jeder ohne großen Aufwand im Internet publizieren, auch nicht-berufliche Kommunikatoren erhalten einen Zugang zur breiten Öffentlichkeit. Das Nadelöhr des klassischen Gatekeepings durch Journalisten und Medienschaffende fällt im Internet weg. Der Zugang wird nun durch einen Mangel an Aufmerksamkeit beim Nutzer eingeschränkt. Diese wenden sich in ihrer Suche nach Orientierung im Internet den Suchmaschinen zu und erwarten von diesen eine ähnliche Form der Qualitätskontrolle wie von Journalisten (Neuberger 2003: 135). Vergleicht man daher die Funktion der Suchmaschinen im Internet mit den Gatekeepern der klassischen Medien in Rundfunk und Presse, so nehmen Suchmaschinen im Netz eine ähnlich zentrale Rolle ein wie der Journalisten.

Während die Kommunikatorforschung die Rolle der Journalisten, deren Arbeitsweisen und Beeinflussbarkeit bei der Nachrichtenauswahl und -gestaltung bereits thematisiert hat (vgl. hierzu insbesondere die Arbeiten von Blöbaum 1994, Requate 1995, Weaver 1998, White 1964, Shoemaker 1991, Galtung/Ruge 1965, Eilders 1997, Kepplinger 2001, Noelle-Neumann/Mathes 1987), fehlen entsprechende empirische Untersuchungen im Hinblick auf die für die Herausbildung einer öffentlichen Meinung so wichtige journalistische Performanz bei Suchmaschinen.

Die Automatisierung von Selektionsprozessen¹¹ wirft die Frage auf, ob menschliches Gatekeeping, klassischerweise eine Leistung von Journalisten, ersetzbar ist. Die Einordnung von Informationen erhält in einem Kommunikationssystem der Informationsüberproduktion besondere Bedeutung; komplexe Sachverhalte müssen mit Hilfe angemessener Darstellungsformen durchschaubar und verständlich gemacht werden (Weischenberg 1985: 197). Journalismus bleibt demnach notwendig, soweit er Informationsqualität garantieren kann, d.h. soweit er im Internet orientiert und über die Qualität fremder Angebote urteilt.

Suchmaschinen sind keine neutralen Roboter, die das öffentliche Geschehen dokumentieren. Wie bei Machill/Neuberger/Schindler (2002) und bei Machill/Neuberger/Schweiger/Wirth (2003) deutlich wurde, gibt es sowohl

10 Ein Weblog oder Blog sind periodische Publikationen von Privatpersonen im Netz, die persönliche Gedanken und Weberlebnisse inklusive Links chronologisch geordnet zum Inhalt haben (vgl. z.B. <http://marketingterms.com/dictionary/blog/> oder <http://de.wikipedia.org/wiki/weblog>).

11 Zentrale Redaktionsaufgaben wie Auswahl, Kategorisierung und Gewichtung der Nachrichten werden bei Google ausschließlich von auf mathematischen Algorithmen basierenden Gruppierungstechnologien erledigt.

interne als auch externe Verfahren, die Ergebnislisten von Suchmaschinen zu manipulieren. Suchmaschinen sind auf der einen Seite programmierte Werkzeuge; hinter ihnen stehen also ökonomische und möglicherweise publizistische Interessen, die es zu durchleuchten gilt. Auf der anderen Seite gibt es zahlreiche Akteure, allen voran die Website-Anbieter, die die scheinbar neutralen Werkzeuge für ihre Zwecke gewinnen wollen: Aufmerksamkeit für die eigenen Internet-Inhalte zu generieren.¹² Beides legt nahe, dass Suchmaschinen als neue „Akteure“ in die Arenen der Öffentlichkeit treten. Suchmaschinen können dabei genauso Manipulationen erliegen wie Journalisten.

Nutzt man an dieser Stelle medienökonomische Ansätze, um zu untersuchen, wie Inhalte ins Mediensystem gelangen, wird evident: Für Werbung müssen Absender bezahlen, um in genau gekennzeichneten Programmflächen (Werbeblöcken, Anzeigenseiten) platziert zu werden. Der Bereich Public Relation lanciert dagegen gezielt Informationen im Rahmen von Organisationskommunikation, die ohne Zahlungsfluss in den redaktionellen Teil der Massenmedien aufgenommen werden. Beide Kommunikationsarten sind Teile des Marketing-Mixes, nämlich der Kommunikations- bzw. Informationspolitik von Organisationen.

Aufgrund von Deregulierung kommt es zu einer Verschiebung der Funktionsabgrenzungen von Journalismus, Public Relations sowie Werbung und damit zu einer Ökonomisierung der Medienindustrie, d.h. Entscheidungsprämissen und Ressourcenallokation werden vor allem ökonomischen Gesichtspunkten unterworfen. Da zudem der Werbemarkt starken Schwankungen unterliegt, also die primäre Einnahmequelle bei hohen Fixkosten unsicher kalkulierbar ist, scheinen die Möglichkeiten der Industrie, auch in den redaktionellen Teil vordringen zu können, zu wachsen. Andererseits sucht aber auch die werbetreibende Industrie nach einer immer prominenteren Platzierung in den Medieninhalten, die der Werbeblock nicht bieten kann. In einem Supplement zum Deutschen Werbekongress 2002 heißt es: „Weil die Konsumenten immer weniger auf Werbeinhalte geben, müssen wir ins Programm“ (Diekhof 2003: 7).

Das Intereffikationsmodell modelliert „eine komplexe Gesamtbeziehung von gegenseitigen Einflüssen, Anpassungen, Erwartungen und Orientierung“ (Bentele 2003: 65) zwischen Journalismus und PR. Danach regt PR durch „Induktionsleistungen“ die Generierung und das Timing von Themen bei Journalisten an (Issue Building, Agenda Setting). Journalisten können durch Adaptionsleistungen (Selektionen, Eigenrecherche, Platzierung und Gewichtung), also durch journalistische Routinen, über Art und Umfang einer derartig induzierten Berichterstattung entscheiden. Von vornherein nicht als

12 Die Aufmerksamkeitsökonomie geht davon aus, dass Aufmerksamkeit zur knappen – und damit wertvollen – Ressource wird, wenn Information im Überfluss vorhanden ist. Darüber hinaus streben Informationssuchende danach, ihre Suchkosten zu minimieren, was beim Vertrauensgut Information nur eingeschränkt möglich ist (vgl. hierzu Franck 1998, Goldhaber 1997, Zerdick et al. 2001: 36–47).

„Symmetriemodell“ angelegt (Bentele 1999: 182), sondern als deskriptives Modell mit möglichen Ausschlägen in beide Richtungen, wird deutlich, dass – durch die Medienkrise und Ökonomisierung mitverursacht – gegenwärtig eine Verschiebung zugunsten des PR-Sektors festzustellen ist. Werden in Redaktionen Personalkosten eingespart, erhöht sich die Chance für Akteure aus dem Bereich der PR, Inhalte direkt, d.h. ohne weitergehende Prüfung lancieren, bzw. durch Medienkooperationen einen großen Teil der redaktionellen Inhalte einnehmen zu können.

Für den Rezipienten, der traditionell in journalistischen Texten eine gegenrecherchierte Vertrauensbasis erwartet, müssen diese Tendenzen zwangsläufig zu einem Vertrauensverlust führen. Einen ähnlichen Vertrauensverlust kann man im Bereich der Suchmaschinen feststellen. Suchmaschinen sind nur erfolgreich, wenn sie als besonders zuverlässig gelten. Während insbesondere die Suchmaschine Google zunächst das Image einer puristischen, nicht-kommerziellen und wegen der Page Rank-Suchtechnologie vor allem „guten“ Suchmaschine (Jungbluth 2003) entwickeln konnte, häuften sich im Sommer 2003 kritische Berichte, die u.a. auf die zunehmende Manipulation von Ergebnislisten hinwiesen.¹³ Die Gatekeeper in der realen Welt und im Cyberspace stehen also gleichermaßen unter Druck.

Bei der Auswahl der Quellen und deren Anzahl bewegen sich alle Suchmaschinenbetreiber zwischen den Ansprüchen an Qualität und Vollständigkeit. Selbst wenn bei der Quellenauswahl eine relative Vollständigkeit erreicht wird (z.B. alle Tageszeitungen aus Deutschland), so bedeutet dies noch nicht, dass sich mit Hilfe der entsprechenden Suchmaschine eine vollständige Presseschau erstellen lässt. Die Online-Angebote der Tageszeitungen enthalten oft nicht die vollständige Print-Ausgabe. In vielen Fällen ist der Zugriff auf die Online-Angebote nur durch eine Registrierung oder Bezahlung möglich (z.B. beim Online-Angebot der *Welt*).

Da die Zahl der registrierungspflichtigen Angebote stark zugenommen hat, haben einige Suchmaschinen Vereinbarungen mit entsprechenden Nachrichtenseiten getroffen, um deren Inhalte indexieren und ihren Nutzern direkt zugänglich machen zu können. So kann beispielsweise über die englischsprachige Seite von Google News auf das Angebot der *New York Times* zurückgegriffen werden, während man sich bei einem direkten Besuch der *Times*-Website registrieren muss, um die Artikel einsehen zu können.

Während die Nachrichten-Websites zudem keinen vollständigen Zugriff auf die Inhalte paralleler Print-Ausgaben erlauben, bieten sie dafür oftmals exklusive Inhalte an, die nicht in der Print-Ausgabe erscheinen. Als Beispiel sei hier *Spiegel Online* genannt, ein Angebot, das nur zu einem kleinen Teil der gedruckten Ausgabe des *Spiegel* entspricht.

13 Vgl. beispielsweise: Borchers 2003, Finkelstein 2003, McHugh 2003, Rötzer 2002.

Die Suchmaschinenbetreiber machen nur selten genaue Angaben zu den erschlossenen Nachrichtenquellen. Die Quellenlisten werden von den meisten Anbietern als Firmengeheimnis gehütet, allenfalls sind Auskünfte über die Quellenmenge zu bekommen. So gibt Google an, international mehr als 4.500 Quellen zu durchsuchen, für das deutschsprachige Angebot sind es mehr als 700. AltaVista nennt 3.000 Quellen weltweit, zur Zahl der deutschsprachigen Quellen werden keine Angaben gemacht.

Eine lobenswerte Ausnahme ist die deutsche Suchmaschine Paperball, die ihre Quellen genau auflistet. Es sind etwa 100 an der Zahl, in erster Linie beschränkt auf Angebote von deutschsprachigen Tages- und Wochenzeitungen und eine Handvoll internationaler Publikationen. Yahoo bietet ein Verzeichnis seiner Partner an, die Nachrichten zuliefern. Dies sind in erster Linie die großen Nachrichtenagenturen.

Die Wahl der Nachrichtenquellen, aus denen der Nachrichtenüberblick eines Angebotes zusammengestellt wird, erfolgt durch redaktionelle Auswahl seitens der Suchmaschinen. Hierbei entscheiden persönliche Vorlieben oder strategische Überlegungen des entscheidenden Redakteurs oder Teams sowie die Firmenpolitik, welche Quellen genutzt werden. Finanzielle oder sonstige Vereinbarungen zwischen den Nachrichtensuchmaschinen und den Anbietern der Nachrichten bestehen nicht. Einzige Ausnahme ist Paperball. Dort hat man die Verwendung einer Meldung der Netzeitung als Topmeldung vereinbart.

Kriterien für eine Auswahl können sein:

- regionale Fokussierung,
- nationale oder internationale Reputation,
- besondere Themenschwerpunkte,
- Anmeldung durch Content-Anbieter.

Kriterien für einen Ausschluss können sein:

- technische Probleme, z.B. beim Spidern/Crawlen¹⁴ der Seiten,
- Entscheidung der Geschäftsleitung oder der Techniker der Publikation (nicht der News-Suchmaschine), den Suchmaschinen-Robots technisch den Zugriff zu verwehren, beispielsweise über die Steuerdatei robots.txt,
- formales oder gar juristisches Verbot des Online-Services hinsichtlich der Aufnahme in die Nachrichtensuchmaschine.

Mit der zunehmende Differenzierung und Arbeitsteilung in den nationalen und internationalen Nachrichtensystemen und dem Anwachsen der interna-

14 Spidern oder Crawlen ist das technische Erfassen und Einlesen von HTML-Seiten in den Index einer Suchmaschine mittels automatischen Programms.

tionalen Berichterstattung als Folge der zunehmenden weltweiten politischen und ökonomischen Verflechtung steigt die Anzahl der Verarbeitungsstufen kontinuierlich an (Schulz 1999: 327). Dieser Entwicklung wirkt die Tendenz der Medien entgegen, Agenturmeldungen oft unverändert oder lediglich vom Ende her gekürzt zu übernehmen (vgl. Hagen 1995). Die Nachrichtenkette vom Ereignis zum Rezipienten ist damit unterschiedlich lang.

Um in dieser Nachrichtenkette die Qualität der Berichterstattung wahren zu können, ist die Transparenz der Nachrichtenquellen oberste Priorität (vgl. Ruß-Mohl, 1994: 94). Sie ist unverzichtbar, sobald der Journalist das Geschehen nicht selbst beobachtet hat (Bruckenberg/Mayr 2004: 167, Schwiesau/Ohler 2003: 83). „Wohlinformierte Kreise“ sind dabei „[g]rundsätzlich problematisch“ (Weischenberg 2001: 139) und sollten – wo möglich – von einer anderen Quelle bestätigt werden.

Aufgrund ihrer zunehmenden Bedeutung hat die Quellentransparenz neben Sachlichkeit, Ausgewogenheit, Fairness, Vollständigkeit, sowie Distanz gegenüber eigenen und fremden Interessen ihren Platz in der Neudefinition journalistischer Objektivität eingenommen (Bondolfi 1999). Massenmedien sind in einer modernen Demokratie Forum öffentlicher Meinungsbildung. Die Offenlegung der Quellen unterstützt den Rezipienten in seiner kritischen Bewertung der ihm zur Verfügung gestellten Nachrichten.

Der Anspruch an transparente Quellen beeinflusst die journalistische Arbeit auf drei Ebenen:

- eigene Recherche
- Verarbeitung von Pressematerial
- Verbreitung von journalistischen Inhalten durch Dritte

Die eigene Recherche unterliegt – wie jede Form der Berichterstattung – dem Anspruch der Sorgfalt. Die Sorgfaltspflicht ist im Pressekodex des Deutschen Presserates verankert¹⁵ und findet ihren rechtlichen Rahmen in den Landespressegesetzen der Bundesländer. Obwohl der deutsche Pressekodex auf die Prüfung der Wahrhaftigkeit der Informationen und deren Trennung von Gerüchten und Vermutungen besteht, bleibt die Transparenz der Quellen außen vor. Die Charta des Schweizer Vereins Qualität im Journalismus¹⁶ hingegen fordert in § 8 Journalisten auf, „fragwürdige Informationen [zu überprüfen], (...) Gegenproben [zu machen] und (...) Transparenz über ihre

15 „Zur Veröffentlichung bestimmte Nachrichten und Informationen in Wort und Bild sind mit der nach den Umständen gebotenen Sorgfalt auf ihren Wahrheitsgehalt zu prüfen. Ihr Sinn darf durch Bearbeitung, Überschrift oder Bildbeschriftung weder entstellt noch verfälscht werden. Dokumente müssen sinngetreu wiedergegeben werden. Unbestätigte Meldungen, Gerüchte und Vermutungen sind als solche erkennbar zu machen.“ (§2 der Publizistischen Grundsätze, beschlossen vom Deutschen Presserat 1973).

16 Die Charta des Vereins Qualität im Journalismus orientiert sich am Kodex „Pflichte und Rechte der Journalistinnen und Journalisten“ (beschlossen vom SVJ 1972).

Quellen [herzustellen], sofern sie nicht gefährdete Informanten schützen müssen.“ Der Offenlegung der Quellen steht im Recherchierjournalismus der Informantenschutz gegenüber. Das Zeugnisverweigerungsrecht deckt dabei alle Personen die „berufsmäßig an die Vorbereitung, Herstellung und Verbreitung“ eines Druckwerks oder einer Rundfunksendung mitwirken und schützt Materialträger und Informanten vor der namentlichen Nennung vor Gericht. Das Zeugnisverweigerungsrecht ist in der bundesdeutschen Strafprozessordnung verankert.

Der Anspruch an transparente Quellen erreicht mit der Zuhilfenahme von Pressemitteilungen seitens des Journalisten eine zweite Ebene. PR-Aktivitäten werden in der Medienberichterstattung bevorzugt als Quellen herangezogen, die Publikationschancen für Pressemitteilungen sind allgemein sehr hoch (vgl. Fröhlich 1992). Fröhlichs Untersuchungen zufolge werden diese Quellen jedoch nur selten durch Journalisten offen gelegt. Die Medien operieren daher unter dem dysfunktionalen Einfluss der Öffentlichkeitsarbeit. Wird die Trennungsnorm bei redaktionell-werblichen Mischformen nicht eingehalten, besteht die Gefahr eines Imageverlustes in Bezug auf die Glaubwürdigkeit dieser Medien (Neuberger 2003: 135). Quellentransparenz ist in diesem Zusammenhang ein Bestandteil der Überparteilichkeit des Journalisten (Schönhagen 1998: 257, 368).

Die dritte Ebene der Quellentransparenz wird durch einen vereinfachten Zugang zur Öffentlichkeit durch das Internet erreicht. Das Fehlen von Qualitätskontrollen und das Anschwellen der Informationsflut – insbesondere durch Weblogs und die ungeprüfte Übernahme von Informationen von Massenmedien durch Massenmedien – lassen die Quellenlage im Internet zunehmend diffus erscheinen (Bondolfi 1999). Die eingeschränkte Aufmerksamkeit auf Seiten der Internet-Nutzer öffnet die Tür für so genannte Advertorials. Hier wird Werbung gezielt als redaktioneller Beitrag gestaltet und ist als kommerzieller Inhalt nicht zweifelsfrei erkennbar¹⁷. Die Medienkompetenz der Rezipienten ist daher im Bereich des Internets von besonderer Bedeutung. Eine Studie der Online News Association belegt, dass Nutzer einen größeren Wert auf die Präsentation der Nachricht legen, während Journalisten als geschulte Medienkonsumenten die Quellentransparenz als wesentlicher einstufen (Neuberger 2003: 135). Es liegt in der Aufgabe der Journalisten, den Rezipienten bei der Hand zu nehmen und ihm durch eine erkennbare Informationsqualität die Unterscheidung von para- und pseudojournalistischen Online-Angeboten zu erleichtern. Qualifizierter Journalismus zeichnet sich auch im Internet durch Vielfalt und Transparenz aus – sowohl im Hinblick auf die Aspekte als auch im Hinblick auf die Quellen (Weischenberg

17 Das japanische Elektronikunternehmen Sony hat in einer 10 Millionen Dollar teuren Kampagne auf mindestens 75 Websites solche Advertorials platziert. Sony wurde in vielen der Texte nicht erwähnt. Erst über Links in Sidebars gelangt der Nutzer auf Sony-Seiten, die über die beschriebenen Produkte informieren.

2002). Gehen diese Ansprüche verloren, droht der Journalismus im Internet unsichtbar zu werden.

2.2.2 Suchmöglichkeiten – Qualitätsmerkmale für Nachrichtensuchmaschinen

Die folgenden Qualitätskriterien für die Nachrichtenrecherche in Suchmaschinen orientieren sich einerseits an den Möglichkeiten der regulären Web-suche, andererseits an den Angeboten professioneller Pressedatenbanken. Die Grundannahme ist, dass Nachrichtensuchmaschinen zur (semi-)professionellen Recherche genutzt werden sollen. Für die Zwecke des Laien, der sich nur kurz einen Überblick zu Pressemeldungen zu einem bestimmten Thema machen möchte, werden keine komplexen Suchfunktionen benötigt. Hier genügen in der Regel einfache Ein- bis Zweiwortanfragen. Für die professionelle Recherche sollten die folgenden Abfragemöglichkeiten zur Verfügung stehen:

- **Quellenauswahl:** Es sollte möglich sein, die Suche sowohl auf eine Quelle als auch auf Quellengruppen zu beschränken. Dabei sollten sinnvolle Gruppen vorgegeben sein, die Erstellung eigener Gruppen sollte jedoch dennoch möglich sein.
- **Regionale Einschränkungen:** Die Suche sollte sich auf bestimmte Gebiete beschränken lassen. Oft sind nur die Ergebnisse lokaler Medien für die Recherche relevant.
- **Ressortbeschränkung:** Die Suche sollte auf einzelne Ressorts beschränkt werden können.
- **Operatoren:** Die Recherche mit Operatoren, vor allem mit den für Presse-recherchen wichtigen Abstandsoperatoren¹⁸, sollte möglich sein. Boole-sche Verknüpfungen bei der Recherche in Volltexten sind nicht ausreichend.
- **Feldbeschränkung:** Die Suche sollte auf den Titel des Artikels und/oder den Anreißer eingeschränkt werden können. Dadurch lässt sich die Präzi-sion der Treffer deutlich erhöhen.
- **Zeitliche Eingrenzung:** Die Suche sollte nicht nur auf die aktuellsten Arti-kel beschränkt werden können, sondern es sollte weiter die Möglichkeit bestehen, einen genauen Zeitraum, in dem die gesuchten Artikel erschie-nen sind, anzugeben.

18 Eine Suche mit Abstandsoperatoren, z.B. Syntax „NEAR“, findet nur Dokumente, in denen die gesuch-ten Begriffe in einer bestimmten Nähe (z.B. im Abstand von maximal zehn Wörtern) zueinander vor-kommen. Solche Suchanfragen finden weniger Dokumente. Dafür vermeiden sie das bei der einfachen UND-Verknüpfung auftretende Problem, dass die gesuchten Begriffe in unterschiedlichen Kontexten stehen.

- Sortiermöglichkeiten: Neben der Sortierung nach Relevanz sollte die Sortierung nach dem Datum möglich sein.
- Cluster-Bildung: Artikel zum gleichen Thema sollten zu thematischen Clustern zusammengefasst werden. Gefundene Artikel sollten auch nach den Ressorts, in denen sie enthalten sind, angezeigt werden können.

Keine der hier untersuchten Nachrichtensuchmaschinen bietet alle diese Funktionen an. Es lässt sich feststellen, dass Abfragemöglichkeiten, die bei der normalen Websuche einzelner Suchmaschinen gegeben sind, für die zugehörige Nachrichtensuche oft nicht vollständig umgesetzt wurden. Erfreulich ist, dass ein Großteil der hier behandelten Suchmaschinen eine Sortierung nach Relevanz oder Datum zulässt.

2.2.3 Überblick über News-Suchmaschinen im deutschsprachigen Raum

Neben eigenständigen Suchmaschinen bieten auch Webverzeichnisse bzw. -portale Elemente der Nachrichtensuche an. Die Funktionen, Suchmöglichkeiten und Philosophien der einzelnen Anbieter unterscheiden sich stark. Die folgende Zusammenfassung soll einen Überblick über die wichtigsten Betreiber im deutschsprachigen Bereich geben.¹⁹

Google News

Google News (news.google.de) existiert in der deutschen Version seit Juli 2003 und befindet sich noch in der Beta-Phase.²⁰ Google verfolgt mit seiner Nachrichtensuche einen neuen Ansatz: Neben der Suchfunktion werden ausgewählte Meldungen vollautomatisch zu einem News-Portal zusammengefasst (Abb.2). Die Nachrichten aus 700 deutschsprachigen Quellen werden in die Rubriken Schlagzeilen, International, Deutschland, Wirtschaft, Wissen/Technik, Sport, Unterhaltung und Gesundheit einsortiert. Zusätzlich werden automatisch Themengruppen erstellt. Beiträge, die sich thematisch ähneln, werden unter „... und ähnliche Artikel“ den jeweiligen Beiträgen zugeordnet. Dabei achtet Google News nach eigenen Angaben auf eine möglichst vielseitige Zusammenstellung der Themengruppen und Ergebnisse und bemüht sich, mehrere Blickwinkel eines Sachverhaltes durch die Präsentation verschiedener Quellen nebeneinander anzubieten.

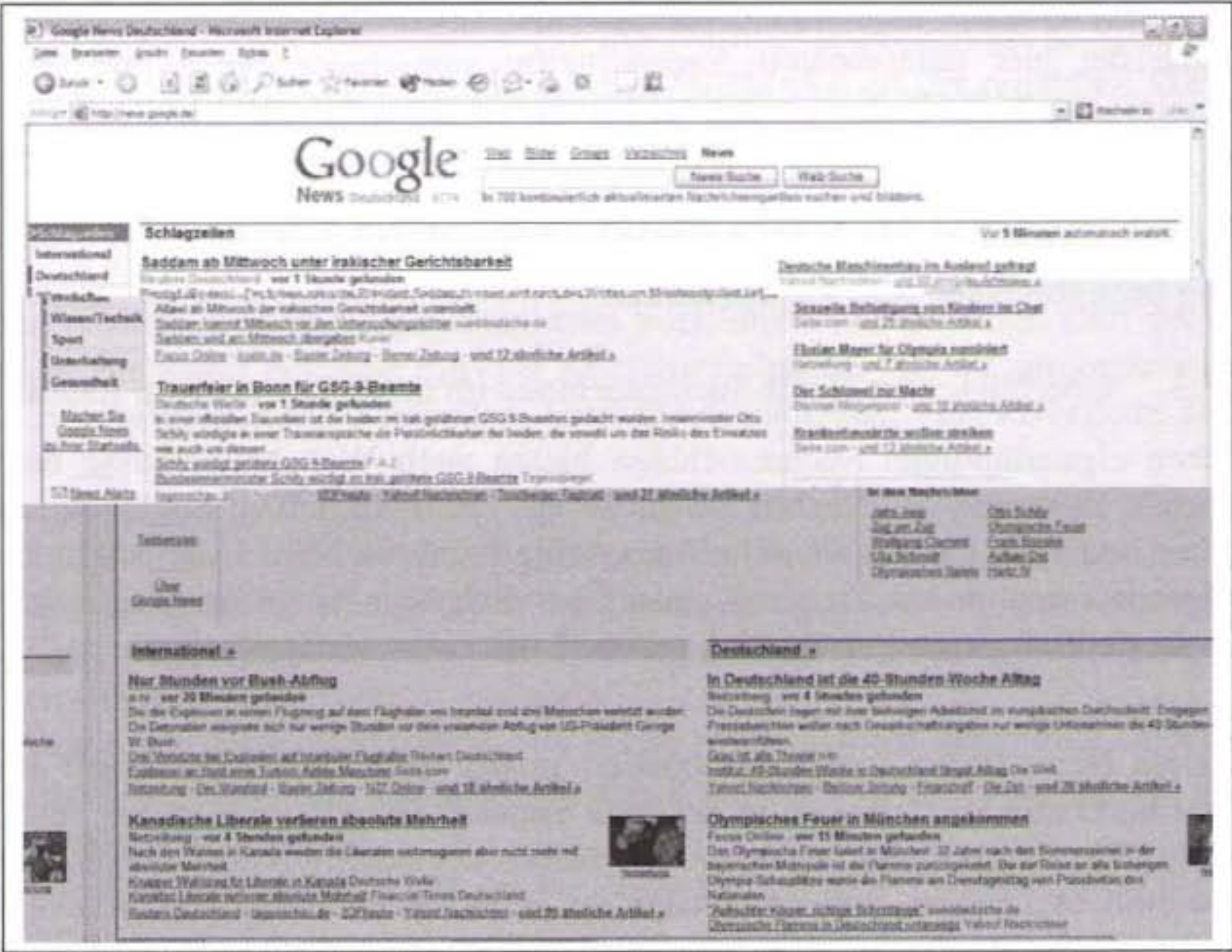
Da die deutsche Google-News-Version sich an ihrem amerikanischen Pendant orientiert, ist zu erwarten, dass in naher Zukunft auch deren erweiterte Funktionen übernommen werden. Dies sind die Möglichkeiten der Einschränkung der Suche auf eine bestimmte Quelle (allerdings nicht auf mehrere) und die Einschränkung nach geographischen Gesichtspunkten, z.B. nur

¹⁹ Die Auswahl ist nicht vollständig und orientiert sich maßgeblich an den in der Kurzstudie untersuchten News-Suchmaschinen.

²⁰ Für gewöhnlich dauern die Beta-Phasen bei Google relativ lange.

Quellen aus einem bestimmten US-Bundesstaat oder einem bestimmten Land. Weiterhin ist es möglich, einen genauen Zeitraum der gesuchten Artikel zu bestimmen.

Abb. 2: Die Startseite der Google-News Deutschland



AltaVista

Bei einer normalen Websuche werden (wenn vorhanden) einige Ergebnisse der Nachrichtensuche vor der jeweiligen Trefferliste eingeblendet. AltaVista sammelt die Nachrichten nach eigenen Angaben aus 3.000 Quellen weltweit. Dabei kann die Suche auf folgende Rubriken beschränkt werden: „Alle Kategorien, Schlagzeilen, Deutschland, Entertainment/Kultur, Finanzen, Lifestyle/Reisen, Märkte/Unternehmen, Sport, Technologie/Medien, Weltweit, Wissenschaft/Gesundheit.“ Zusätzlich kann die Suche nach Datum eingeschränkt werden, wobei entweder nach vorgegebenen Zeiträumen oder nach einer bestimmten Zeitspanne gesucht werden kann. Die Treffer enthalten neben der Überschrift und einem zwei- bis dreizeiligen Abstract einen Hinweis zur Quelle sowie die Information, wann die Meldung gefunden wurde.

Fireball-News und Lycos-News: Paperball

Paperball ist der spezielle Nachrichtensucher für die Suchmaschinen Fireball und Lycos, die beide zur Lycos-Gruppe gehören. Paperball wertet etwa 100 Tageszeitungen aus. Hinzu kommen einige Wochenzeitungen und Zeit-

schriften. Die Nachrichtensuchmaschine unterteilt die Suchergebnisse in die Bereiche „Überregional“, „Regional“, „International“. Die voreingestellte Suche (Default-Suche) erfolgt über einen Zeitraum von zwei Tagen. Wahlweise ist es möglich, die Suche auch auf sieben Tage auszudehnen.

Auf der Startseite ist die jeweilige Top-Meldung der Netzeitung eingeblendet. Diese Einblendung wurde zwischen der Netzeitung und Paperball vereinbart. Über diese Übereinkunft hinaus bestehen keine weiteren Verträge.

Abb. 3: Die News-Suche in Lycos. Die Paperball-Suchmaske findet sich erst im unteren Teil der Seite



Paperball bietet die Möglichkeit, die Suche nur in den aktuell gefundenen Artikeln oder aber im Nachrichtenbestand der letzten sieben Tage durchzuführen. Genauere Datumsbeschränkungen sind nicht möglich. Dafür bietet Paperball eine Einschränkungsmöglichkeit, die sich bei anderen Suchmaschinen für das deutsche Sprachgebiet nicht finden lässt: die regionale Beschränkung der Suche. In der Regionalauswahl sind die Quellen nach Bundesländern angeordnet; ergänzt werden sie von überregionalen und ausländischen Quellen. Für die Suche kann aus dem Quellenspektrum eine beliebig kombinierte Zusammenstellung erstellt werden. Einzelne Quellen können so ausgeschlossen werden. Weiterhin bietet Paperball die bei anderen Suchmaschinen

meist fehlende Möglichkeit, die Suche auf den Titel („title:“), den Anreißer („teaser:“) oder auf den Autor des Artikels („writer:“) einzuschränken.

Fireball bietet zur Aktivierung der News-Suche einen Auswahlknopf („Radiobutton“) unter dem Suchfeld. Die Aktivierung des Buttons „Nachrichten“ in Fireball startet sofort Paperball in eigenständigem Layout. Bei der normalen Suche in Fireball werden daher keine Treffer aus der Nachrichtensuche eingeblendet. Bei Lycos ist Paperball als unabhängige Applikation auf der News-Startseite mit einer Suchmaske in der unteren Hälfte der Seite präsent. Dort findet sich auch die Eingabemaske zur dpa-Archiv-Suche.

Yahoo

Yahoo verfügt über ein eigenes Nachrichtenportal. Aktuelle Meldungen werden auf der Startseite präsentiert. Yahoo-News bieten keine Möglichkeiten der erweiterten Suche. In der Standardeinstellung werden die Suchbegriffe sowohl links als auch rechts trunkiert. Dies erhöht auf der einen Seite die Ergebnismenge, senkt auf der anderen Seite jedoch deren Präzision erheblich. Um präzise Fundstellen zu erhalten, müssen die Begriffe einzeln in Anführungszeichen gesetzt werden.

Unter Yahoo-Nachrichten finden sich die Ressorts „Kurzberichte, Wirtschaft, Politik, High-tech, Wissenschaft, Unterhaltung, Lifestyle, Vermischtes, Sport“ und sogar „Video“. Diese kurzen Nachrichtenvideos werden von Reuters geliefert. Außerdem bedient sich Yahoo der Agenturen und News-Quellen AP, backpage (BMG), ddp, Deutsche Welle, doc rock (infomusic), dpa, expedition zone, Lifeline, n24, News Aktuell, n-tv, Reuters, Shownet, tecchannel, womenweb, zdf.msnbc und zdnet. Pressemeldungen und Ticker-Meldungen werden in der Trefferliste gemeinsam aufgeführt und nach Quellen gekennzeichnet. Die Rubriken werden teilweise im fünfminütigen Rhythmus aktualisiert.²¹

Alltheweb

Die Suchmaschine Alltheweb bietet grundsätzlich keine länderspezifischen Interfaces an. So findet auch bei der Nachrichtensuche keine Spezifizierung statt. Standardmäßig wird stets in allen Quellen weltweit gesucht. Wahlweise erfolgt die Suche in allen verfügbaren Sprachen oder in den vom Nutzer individuell bzw. von Alltheweb in der Standardeinstellung voreingestellten Sprachen. Eine genaue Spracheinschränkung lässt sich in der erweiterten Suche vornehmen. Zudem ist eine lokalisierte Suche nach einem Herkunftsland oder einer bestimmten Stadt möglich. Bei Alltheweb lässt sich die Suche auch nach Rubriken einschränken, die Artikel werden automatisch zugeordnet. Weiterhin lässt sich die Suche nach Datum einschränken, wobei kein genauer Zeitraum angegeben, sondern nur auf festgelegte Zeiträume zurückgegriffen werden kann.

21 Vgl. Yahoo-Hilfeseiten (<http://eur.help.yahoo.com/help/de/news/news-06.html>).

MSN Newsbot

MSN baut derzeit ein englischsprachiges Nachrichtenportal auf. Hinweise auf dem deutschen Portal finden sich jedoch noch nicht. Die Suchergebnisse werden von Inktomi aus dem normalen Webindex geliefert. Die Nachrichtensuche erfolgt nur in den kooperierenden Magazinangeboten *Fokus-Online-News*, *MSN Money*, *Max*, *Amica*, *TV Spielfilm*, *Cinema*, *Tomorrow*, *Fit For Fun*, und *Expedia*. Der Bereich Nachrichten wird ausschließlich von *Focus* bestritten. Der Microsoft-Nachrichtensuchdienst bietet bisher keine erweiterten Suchfunktionen an.

Web.de

Die Nachrichtensuche ist nicht direkt auf der Startseite durchführbar. Erst mit einem Klick auf "Schlagzeilen" gelangt man in den Nachrichtenbereich, wo auf jeder Meldungsseite das Suchfeld für die Nachrichtensuche integriert ist. Der Link „Archiv“ führt zu einer chronologischen Liste der Meldungen der vergangenen Tage. Diese Meldungen werden von dpa geliefert.

AOL

AOL hat keine eigene News-Suche, sondern verwendet ausschließlich Google-Ergebnisse der normalen Websuche. Entsprechend ungeeignet ist die Suche zur Erfassung von Neuigkeiten.

T-Online

Nachrichten sind ein zentraler Bestandteil des Portals. Das DPA Newsfeed wird komplex verwoben präsentiert. Die News Suche hingegen ist über die normale Suche nicht erreichbar. Erst mit einem Klick auf „Profi-Suche“ erreicht man eine Suchbox, auf der die Wahl zwischen Web, News, Bildern und Suchcenter möglich ist (Abb. 4). Die Nachrichtensuche durchsucht jedoch nicht das DPA Newsfeed, sondern ausgewählte Websites. Eine Sortierung nach Datum findet nicht statt.

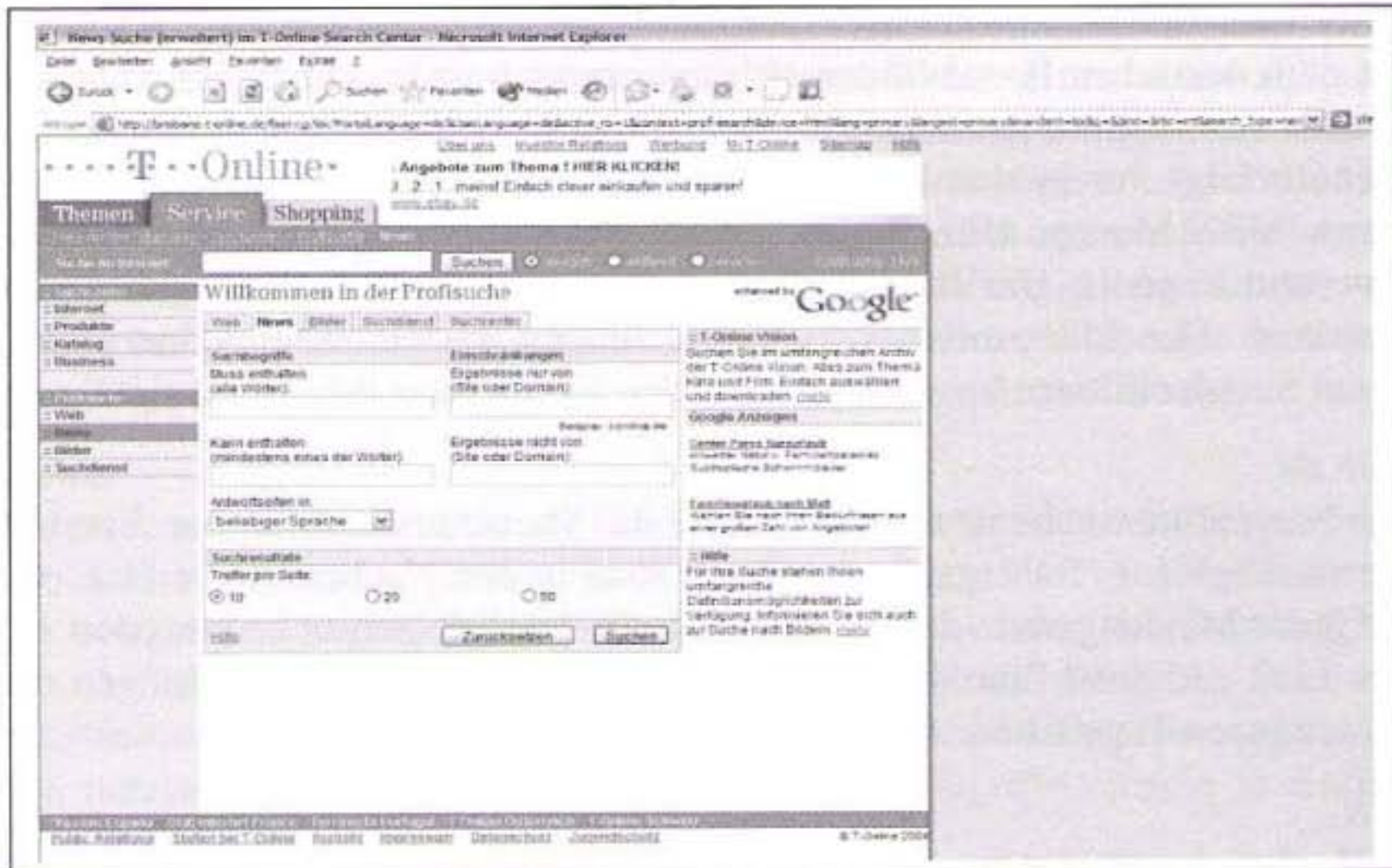
Metager

Nachrichtenergebnisse werden bei Metager nicht besonders berücksichtigt. Es werden nur reguläre Suchmaschinen eingebunden. Die Metager-Ergebnisliste ist – anders als bei anderen Suchdiensten – deutlich länger als zehn Treffer. Teilweise werden bis zu 200 von ihnen angezeigt.

Paperazzi

Paperazzi ist eine reine Nachrichtensuchmaschine. Abgesehen von der direkten Konkurrentin Paperball ist sie die einzige professionelle Nachrichtensuchmaschine unter den deutschsprachigen Suchmaschinen.

Abb. 4: Die T-Online-News-Suche ist nur über die Profisuche erreichbar



2.2.4 Präsentationsformen

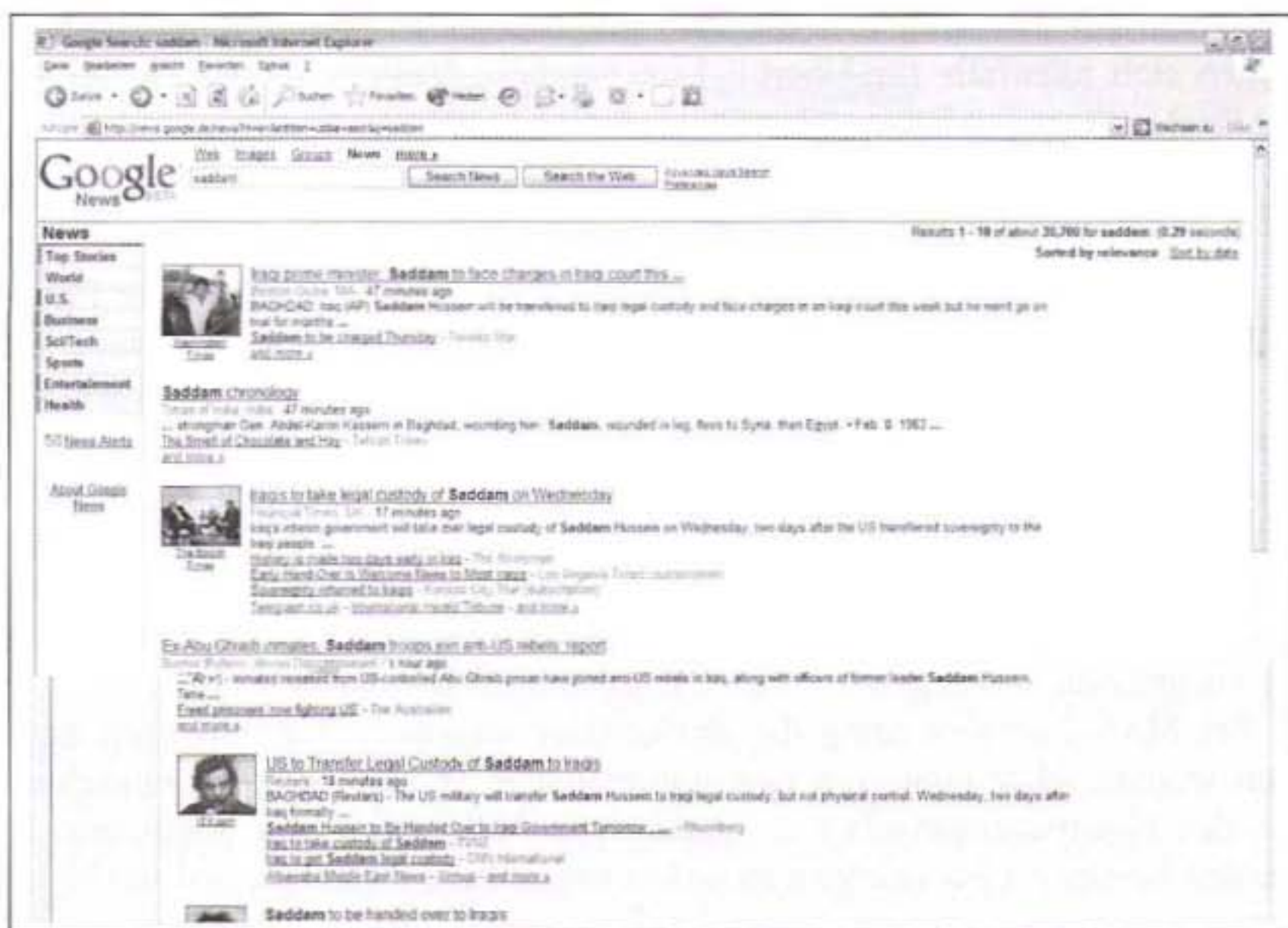
Nachrichtensuchmaschinen dienen nicht nur der eigentlichen Suche nach Presseartikeln, sondern auch dem Überblick über aktuelle Ereignisse. Wie bei der Erschließung regulärer Internet-Quellen bestehen bei Nachrichtensuchern die zwei Zugriffsebenen „Searching“ und „Browsing“. Beim „Searching“ wird gezielt ein Dokument bzw. eine Menge von Dokumenten als Ergebnis gewünscht. Hierbei handelt es sich also um eine systematische Aktion des Nutzers. Beim „Browsing“ folgt der Nutzer – ausgehend von einem Startdokument – den verweisenden Links. Um beiden Erschließungsansätzen gerecht zu werden, bieten einige Suchmaschinen neben den Suchseiten auch Übersichtsseiten, die aktuelle Nachrichten enthalten.

So besteht die Startseite des Nachrichtenangebots bei einigen Anbietern (z.B. Paperball, AltaVista) aus Links auf aktuelle Nachrichten. Teilweise werden dazu kurze Zusammenfassungen angeboten. Diese Version der Startseite findet sich vor allem bei den schon länger am Markt agierenden Anbietern.

Die zweite Variante betont den „Browsing“-Ansatz und präsentiert eine ausführliche Übersichtsseite mit aktuellen Nachrichten der verschiedenen Rubriken. Handelt es sich um eine Suchmaschine, die nicht selber nach Nachrichteninhalten crawlt, sondern ihre Ergebnisse zukauf, erfolgt die Einordnung auf der Basis der mitgelieferten Metadaten. Gewinnt die Suchmaschine ihre Inhalte durch einen eigenen „Crawling“-Prozess, so müssen die Nachrichten zur Erstellung der Übersichtsseite nach inhaltlichen Merkmalen gruppiert und den Ressorts zugeordnet werden. Durch diese Gruppierung von

Nachrichten zum gleichen Thema wird verhindert, dass inhaltlich gleiche Meldungen mehrfach auf der Übersichtsseite erscheinen.²² Die Meldung sollte einmal inklusive Anreißer (und einem Bild) gelistet und um Verweise auf weitere Meldungen zum gleichen Thema ergänzt werden. Diese Methode wird für Google News eingesetzt. Diese Suchmaschine war der Vorreiter für die automatische Zusammenstellung der Übersichtsseiten. Der neue MSN Newsbot verwendet ein ähnliches Verfahren. Daher ist anzunehmen, dass die bestehenden wie auch eventuell neuen Nachrichtensuchmaschinen in Zukunft verstärkt den Anforderungen des Browsing genügen werden und automatisch zusammengestellte Übersichtsseiten anbieten.

Abb. 5: Die Ergebnisliste einer Suchanfrage bei news.google.de



2.2.5 Archivfunktionen

Nachrichtensuchmaschinen reichen in ihren Archiven nicht weiter zurück als die Nachrichten-Sites selbst. In der Regel sind die Archive von Nachrichten-

22 Da diese Einordnung automatisch vorgenommen wird, sind die betreffenden Systeme noch fehleranfällig. So erscheint die Meldung über einen Angriff von tsetschenischen Rebellen in der russischen Teilrepublik Inguschetien auf news.google.de am 22. Juni 2004 unter den Überschriften „Innenminister bei Rebellenangriff getötet“ sowie „Mindestens 16 Tote bei schweren Unruhen“ gleich doppelt. Auf der gleichen Seite wurden im Ressort Unterhaltung die Meldungen „Hamburger Filmförderung halbiert“ und „Hamburg halbiert die Filmförderung“ direkt untereinander gelistet (vgl. hierzu auch: 2.3.3 Fast identische Meldungen).

angeboten vier Wochen lang kostenlos verfügbar. Daran orientieren sich auch die Suchmaschinen (Schiff 2003). Artikel, die vor Ablauf dieser Zeitspanne nicht mehr kostenlos zugänglich sind, fallen ebenfalls aus den Indizes der Suchmaschinen heraus. Gleiches gilt für registrierungspflichtige Angebote. Deren Inhalte können durch Suchmaschinen nicht erschlossen werden, da sie in den Bereich des Invisible Webs²³ fallen. Die Suchmaschinen stoßen bei vorgeschalteten Formularen an Zugangsgrenzen und können nicht bis zu den Inhalten vordringen (vgl. Spencer 2002).

Die Qualität der Archivsuche in Nachrichtensuchmaschinen ist dementsprechend unbefriedigend, da schon aufgrund der Quellenproblematik keine Vollständigkeit erreicht werden kann. Diese Unvollständigkeit wird durch die unterschiedlichen Zugriffszeiten der kostenlosen Archive verstärkt. An der Brauchbarkeit der Archive muss also grundsätzlich gezweifelt werden: Sie eignen sich allenfalls für Überblicksrecherchen. Vollständigkeit ist nur bei der Recherche in professionellen Pressedatenbanken zu erreichen.

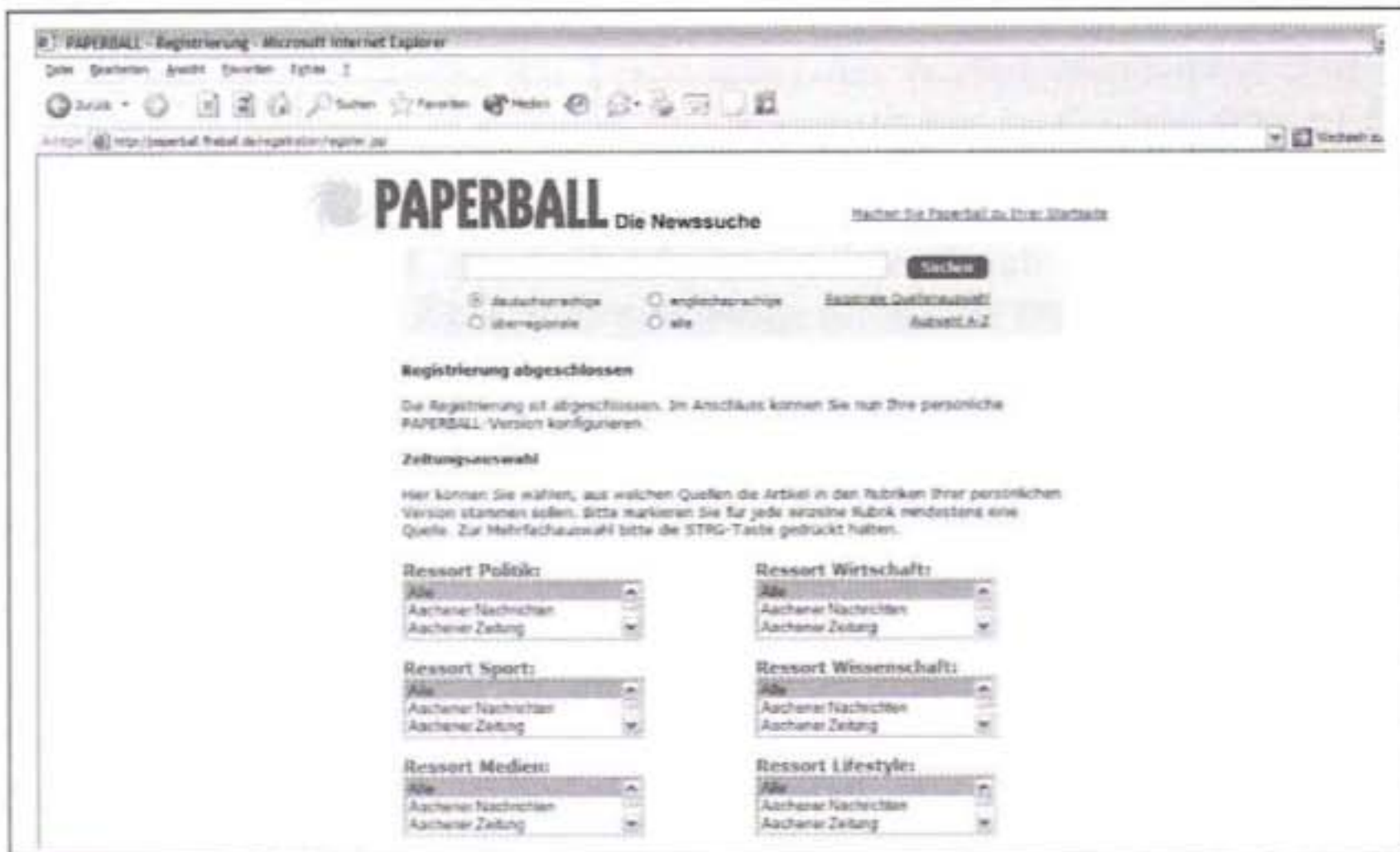
2.2.6 Personalisierung

Paperball und MSN Newsbot bieten dem Nutzer die Möglichkeit der Personalisierung. So ist es bei Paperball möglich, individuell Nachrichtenquellen zu einer „persönlichen Zeitung“ zusammenzustellen. Dabei können die Quellen nach einzelnen Rubriken ausgewählt werden. So lässt sich beispielsweise der Politikteil einer überregionalen Tageszeitung mit dem Wirtschaftsteil einer spezialisierten Wirtschaftszeitung kombinieren. Für jede Rubrik lassen sich eine oder mehrere Quellen auswählen. Die Einstellungen werden gespeichert. Die „persönliche Zeitung“ kann dann im Ganzen angesehen oder für die gezielte Suche genutzt werden.

Bei MSN Newsbot muss die Personalisierung nicht manuell vorgenommen werden. Hier lässt sich eine automatische Personalisierung einrichten, die das Nachrichtenangebot am bisherigen Nutzerverhalten ausrichtet: Es werden bevorzugt Nachrichten zu bisher angesehenen Themen und aus bisher benutzten Quellen zusammengestellt. Diese Funktion ist allerdings nur für Microsoft-Passport-Nutzer zugänglich.

23 Unter Invisible Web werden jene Bereiche des WWW verstanden, die nicht frei zugänglich sind, d.h. es handelt sich um Daten, die durch Passwörter oder ähnliche Zugangsbeschränkungen geschützt werden.

Abb. 6: Die Personalisierungsoptionen in der Paperball-Nachrichtensuche



2.2.7 Monitoring

Während alle bisher behandelten Punkte den so genannten „Pull-Ansatz“ – das heißt die aktive Recherche des Nutzers durch Eingabe von Suchbegriffen – zugrunde legen, basiert Monitoring auf dem „Push-Ansatz“. Hierbei werden aufgrund eingerichteter Profile neue Meldungen, die der Suchanfrage entsprechen, ausgegeben. Einerseits kann dies beim Besuch der Nachrichtensuchmaschine und der Identifikation des Nutzers geschehen, andererseits aber auch durch E-Mail-Benachrichtigungen. Ein solches Verfahren ist als „selective dissemination of information“ (SDI) bekannt (Chu 2003, 83).

Bisher bieten Paperball und Google News eine solche Funktion. Bei Google lassen sich unter „News Alerts“ entsprechende Profile einrichten. Die Ergebnisse werden entweder bei jeder Aktualisierung oder einmal am Tag per E-Mail verschickt. Auch Paperball bietet eine ähnliche Funktion an. Sie ist für das Monitoring jedoch oft nicht ausreichend, da Paperball sein Quellspektrum bewusst auf die Angebote der Tages- und Wochenpresse beschränkt.

2.3 Probleme bei der Nachrichtensuche

Nachrichtensuchmaschinen, die ihre Daten mittels automatischer Robots indexieren, greifen (derzeit noch) nicht auf strukturierte Datenströme wie etwa XML zu, die eine automatische Erkennung und Zuordnung der Inhalte zulassen. XML definiert genau, welcher Teil eines Dokumentes Überschrift oder Abstract ist. Bei HTML fehlt diese Möglichkeit, Seitenelemente sind

nicht automatisch inhaltlich und strukturell zuzuordnen. Daher müssen Nachrichtensuchmaschinen ihre Daten selbst aufbereiten.

2.3.1 Erkennen der tatsächlichen Nachricht

Eine Webseite mit einem Nachrichtenartikel besteht aus einer Reihe von Elementen, die für das Funktionieren der Seite im Rahmen eines Online-Nachrichtenangebotes zwingend notwendig sind, jedoch nicht in Verbindung zum Inhalt der Nachricht stehen. So enthält jede Seite eine Hauptnavigation, Hinweise auf weitere Angebote, Links zu verwandten Themen oder anderen aktuellen Nachrichten sowie Links zu Partnerseiten und externen Informationsquellen. Dazu kommen Werbeanzeigen in Bild- oder Textform. All diese Informationen sind für die eigentliche Meldung nicht nur irrelevant, sie bergen auch die Gefahr, dass einem Text Stichwörter zugeordnet werden, die nicht in Verbindung mit dem Artikel stehen. Dieser unerwünschte Kontext muss gefiltert werden.

Diese Filterung geschieht, indem jeder Seitentyp einer Meldung auf seine typischen Merkmale hin untersucht wird. Der Pfad auf dem Webserver oder eine typische URL können helfen, die Dokumente zu identifizieren, die die Nachricht enthalten. Im Dokument selbst muss nach typischen Elementen der Seitenauszeichnung gefahndet werden, um die für die Meldung wichtigen Bestandteile isoliert erfassen zu können. Üblicherweise sucht man nach typischen HTML-Code-Schnipseln, die dank der automatischen Generierung immer identisch sind.²⁴ Dieser Auswahlprozess muss manuell durchgeführt und maschinell überwacht werden. Ein Wechsel des Layouts bzw. Seitenaufbaus eines Nachrichtenangebotes, der anhand von Überwachungsprogrammen schnell erkannt wird, zieht immer auch eine manuelle Nachbesserung der Konfiguration der Identifizierungs-Software nach sich. Diese Prozedur kann mit geeigneter Software zumindest halbautomatisiert werden.

2.3.2 Aktualität von Meldungen im WWW – Bestimmung des Nachrichtendatums

Datum und Uhrzeit einer Meldung sind in den seltensten Fällen aus dem HTML-Dokument sofort und ohne Umwege ersichtlich. Häufig wird in den Seiten das jeweils aktuelle Datum des Aufrufs einer Seite angegeben. Dementsprechend ist es eine schwierige Aufgabe, zu bestimmen, wann eine Meldung tatsächlich publiziert wurde.

Für die Nachrichtensuchmaschinen ist die Bestimmung des Nachrichtendatums jedoch von großer Bedeutung, da sie die Suchergebnisse automatisch ordnen. Der Aspekt der Aktualität ist für das Ranking der einzelnen Ergebnisse sehr wichtig. So würde eine Suchmaschine, die zwar relevante aber

²⁴ Die Meldungen werden anhand eines einheitlichen Templates, also Layout-Masters erstellt.

veraltete Meldungen an den ersten Ergebnisplätzen anzeigt, bald an Nutzern verlieren (vgl. van Eimeren/Gerhard/Frees 2003: 348).

Die meisten Methoden zur Feststellung des Nachrichtendatums sind mit einer hohen Fehlerquote behaftet und können leicht in die Irre führen. Folgende Möglichkeiten gibt es, eine Datumsangabe anzuzeigen:

Redaktionelles Datum

Eine Möglichkeit der Datumsbestimmung ist das „Ursprungsdatum“. Darunter versteht man Datum und Uhrzeit einer Meldung, die manuell oder automatisch zusammen mit der Meldung vom ursprünglichen Autor (z.B. dpa) ausgegeben werden. Dieser Zeitwert ist Teil der Originalmeldung, die vom Presse-Service verschickt wird. Die spätere Darstellung der Zeit im letztendlich veröffentlichten Artikel kann, muss aber nicht diesen Wert übernehmen. Das Ursprungsdatum ist auch der Basiswert, der für die Aktualität einer Meldung von Bedeutung ist. Dieser Wert als Zeit-/Datumsangabe kann auf der Website an beliebiger Stelle auftauchen. Jedoch machen nur die wenigsten Angebote hiervon Gebrauch.

Eine zweite Form des redaktionellen Datums ist die „Erstpublikation“. Hierunter wird das Datum und die Uhrzeit der Erstpublikation der Ticker-Meldung innerhalb des Nachrichtenportals verstanden. Zu dieser Uhrzeit wird die Meldung erstmals einer breiten Öffentlichkeit zur Verfügung gestellt. Dieser Zeitpunkt liegt nach der Veröffentlichung durch die Nachrichtenagentur, da zwischen dem Publizieren der Ursprungsmeldung durch den Lieferanten der Rohnachrichten (etwa dpa oder Reuters) und der automatischen oder manuellen Übernahme der Meldung in das Layout der jeweiligen Publikation Zeit verstreicht, im Normalfall einige Minuten.

Technisches Datum

Das technische Datum ermittelt die Aktualität eines Beitrages auf der Basis von Informationen des Dateisystems. Eine erste Form ist das Datum der *erstmaligen Erstellung*. Dieses Datum markiert den Zeitpunkt, an dem eine HTML-Seite erstmalig auf dem Server des Nachrichtenverteilers erstellt wurde; also den Zeitpunkt der erstmaligen physischen Erstellung der Seite. Dieser Wert wird vom Webserver als Erstellungsdatum an den Client gesendet, sofern es sich nicht um ein dynamisch erstelltes Dokument handelt.

Eine weitere Form der Aktualitätsbestimmung ist das Datum der *letzten Aktualisierung*. Sie bezeichnet Datum und Uhrzeit der letzten technischen Aktualisierung der Webseite. Dies geschieht etwa, wenn sich Links im Kontext geändert haben, neue Werbung eingespielt oder, ein neues Layout eingeführt wurde, ohne dass sich der Text der Meldung selber geändert hätte. Bei der Aktualitätsbestimmung über dieses Datum kann eine Meldung aktueller erscheinen, als sie tatsächlich ist.

Auch die Server-Variable (Zeit im Http-Header der Webseite beim Aufruf) kann zur Aktualitätsbestimmung einer Meldung herangezogen werden.

Eine Server-Variable ist ein Datum, das vollautomatisch vom Webserver auf der HTML-Seite eingestanzt wird und sich dem Zeitverlauf entsprechend jede Sekunde ändert. Diese Zeit stellt somit die Server-Zeit des Seitenaufrufes fest.

Bei Publikationen mit Content-Management-Systemen (CMS) kann das jeweilige Generierungsdatum im CMS zur Bestimmung von Aktualität herangezogen werden. Das Generierungsdatum der Meldung im Content-Management-System des Anbieters kann die Zeit der erstmaligen Speicherung der Nachricht im System sein oder die der ersten (oder auch letztmaligen) Publikation bezeichnen.

Datum und Uhrzeit in Meta Tags

Das entsprechende Feld für Zeitangaben in Meta Tags kann durch das verwendete CMS (und damit manuell oder automatisch) oder durch Server-Variablen gefüllt werden. Entsprechend ist an dieser Stelle keine Zuordnung auf eine bestimmte Zeit möglich. Datum und Zeit können beispielsweise manuell zurückgesetzt sein und unbrauchbare Werte wie `<meta name="generated" content="01.01.1970 - 01:00" />` anzeigen.

Änderungsdatum

Änderungsdatum und erstmaliges Generierungsdatum lassen sich nicht voneinander unterscheiden. Der Zeitstempel auf einer HTML-Seite kann durch das System oder einen redaktionellen Eingriff geändert worden sein, ohne dass das ursprüngliche Datum ersichtlich wäre.

2.3.3 Fast identische Meldungen

Ticker-Meldungen werden oft unter großem Zeitdruck geschrieben. Entsprechend schleichen sich ab und zu Tippfehler ein. Häufig treffen auch ergänzende Informationen zu einem Sachverhalt ein und die Meldung wird umgeschrieben. Die Nachrichtenagenturen teilen jeder neu herausgegebenen Meldung eine neue Identität zu. Sie stellen Mechanismen zur Verfügung, um aktualisierte Meldungen und Doppler mit Meldungen weitgehend identischen Inhalts, aber unterschiedlicher Version zu erkennen. Die Fehlerquote ist bei diesem Prozess aber noch sehr hoch, so dass sich auf dem Server des News-Anbieters immer wieder sehr ähnliche Meldungen befinden und von den News-Suchmaschinen auch prompt aufgenommen werden.

2.3.4 Ranking von Nachrichtenmeldungen

Genau wie bei der normalen Websuche machen die Suchdienste nur ungenaue Angaben über ihre Ranking-Kriterien.²⁵ Man kann aber einige Faktoren

²⁵ Konkrete Vorschläge, was eine Veröffentlichung der Ranking-Kriterien beinhalten sollte, stellt der „Code of Conduct“ der Bertelsmann Stiftung vor (vgl. Welp 2003: 493 ff.).

erkennen, die bei den einzelnen Anbietern in unterschiedlicher Gewichtung die Sortierung der Treffer beeinflussen.

Datum und Zeit

Datum und Zeit einer Meldung sind wesentliche Faktoren für deren Aktualität. Aus diesem Grund bietet beispielsweise Google die Listung der Treffer entweder nach Aktualität oder Relevanz an. Jedoch ist die automatische Feststellung von Aktualität, wie in 2.3.2 ausgeführt, mit großen Problemen behaftet und noch sehr ungenau.

Stichwörter in bestimmten Dokumententeilen – Feldsuche

Die Feldsuche in Titel, „Subline“ (oft als „Teaser“ verwendet) und Text der Meldung sowie die Gewichtung des Vorkommens von Stichwörtern in diesen ausgewählten Bereichen sind eine klassische Ranking-Methode. Die gemessen an anderen Ranking-Methoden vergleichsweise einfachen Regeln der Gewichtung in Form der Zählung der Stichwörter in den einzelnen Bereichen eines Datensatzes funktionieren bei den ausgewählten Quellen des Datenraumes sehr viel besser als bei den unstrukturierten Dokumenten des gesamten WWW. Die Gefahr, dass die Gewichtung durch Manipulation verfälscht wird (vgl. Karzauninkat 2003, S. 528 ff.), ist praktisch nicht vorhanden. Damit bietet diese Methode einen deutlich solideren Ausgangspunkt als bei der normalen Websuche. Auf technischer Seite problematisch sind lediglich launige und reißerische Titel, die keine automatische Zuordnung von Stichwörtern zum Inhalt der Meldung zulassen. Ein speziell für Suchmaschinen zusätzlich erstellter, sachlicher Titel wäre hier hilfreich.

Zuordnung zu Ressorts

Zumindest bei der Formulierung der Anfrage kann die Suche auf ein oder mehrere Ressorts eingeschränkt werden. Die Zuordnung der Quelldokumente zu bestimmten Ressorts seitens der Nachrichtensuchmaschine erfolgt durch redaktionelle Vorauswahl des Teams und maschinelle Erkennung von entsprechend identifizierten Stichwörtern im Dokument, die die Zugehörigkeit zu dem Ressort bestimmen.

Regionale Gewichtung

Die Einordnung der Quelle in regionale, überregionale oder internationale Kategorien, lässt eine Gewichtung verschiedener Publikationen nach Themengebieten zu. So kann mittels einer Stichwortliste üblicherweise überregional behandelte Themen den überregionalen Quellen zusätzliches Gewicht beim Ranking verliehen werden.

2.3.5 Copyright

In einem Grundsatzurteil hat der Bundesgerichtshof entschieden, dass „Deep Links“ auf im Web veröffentlichte Nachrichten, die unter Umgehung der

Startseite direkt auf die Nachricht verweisen, zulässig sind (BGH 2003). Damit wurde zumindest für die Nachrichtensuchmaschinen eine erste Rechtsgrundlage geschaffen. Ungeklärt ist, in welchem Umfang eine Nachrichtensuchmaschine Daten übernehmen darf, und über welchen Zeitraum die Ergebnisse zugreifbar bleiben dürfen. Viele Online-Publikationen lassen eine Archivsuche nur für registrierte Nutzer zu oder bieten Archivseiten ausschließlich kostenpflichtig an. Insofern könnte ein allzu ausführlicher Abstract eines Treffers auf eine nicht mehr öffentlich zugängliche Seite das Copyright verletzen.

Schon jetzt folgen alle Nachrichtensuchmaschinen dem eventuellen Wunsch einer Publikation, nicht durchsucht zu werden. Die Klärung, ob eventuell ganze Angebote oder Teile davon nicht aufgenommen werden dürfen, erfolgt entweder über die robots.txt, die den Crawlern den Zugriff verwehren kann, oder über persönlichen Kontakt zum Suchmaschinenbetreiber.

3 Kurztest: Vergleich der Suchergebnisse von zehn Suchmaschinen und deren separater Nachrichtensuche (falls vorhanden)

3.1 Testaufbau und Trefferwertung

Untersucht wurden die Suchmaschinen Alta Vista, Google, Yahoo, Fireball, Lycos, T-Online, Web.de, MSN, AOL und Metager, sowie ergänzend im Reaktionstest die unabhängige Nachrichtensuchmaschine Paperazzi. Diese Suchmaschinen wurden hinsichtlich ihrer Reichweite und Markbedeutung in Deutschland ausgewählt (vgl. 2.2.3). Der Test gliedert sich in folgende drei Teilbereiche:

Fünf-Tage-Test

Das Experiment über fünf Tage stellt fest, inwieweit die Nachrichtensuchmaschinen sowohl relevante als auch aktuelle Suchergebnisse zu einem aktuellen Ereignis anbieten. Der Test liefert zusätzlich Daten für die Untersuchung der Treffer nach Verbreitungsgebiet und eventueller Verflechtungen. Die Suchmaschinen wurden über fünf Tage hinweg täglich gegen 11.00 Uhr mit identischen Suchanfragen (Querries) abgefragt. Die verwendeten Suchanfragen lauteten:

- Schill
- Irak Opfer
- Toll Collect Maut
- Simbabwe Commonwealth
- Parlamentswahl Russland

24-Stunden-Test

Die Untersuchung innerhalb von 24 Stunden sollte Aufschluss darüber geben, ab wann im Tagesverlauf die Treffer in den Suchmaschinen auftauchen. Zusätzlich wurden bei allen Suchdiensten mit separater Nachrichtensuche die Ergebnisse derselben mit denen der normalen Websuche verglichen. Hierfür wurden die zehn Suchmaschinen jeweils zehn und 24 Stunden nach einem Ereignis getestet. Die verwendeten Suchanfragen lauteten:

- Explosion Bagdad
- Schill
- Flugpassagierdaten
- Vermittlungsausschuss
- SARS Taiwan

Reaktionstest: Vergleich +1h +3h +5h

Mit dem Reaktionstest wurde die Geschwindigkeit der Suchmaschinen bei der Aufnahme von aktuellen Ereignissen in den Nachrichtenindex untersucht. Neun Suchmaschinen wurden eine, drei und fünf Stunden nach einem von Reuters gemeldeten Zeitpunkt eines Ereignisses auf ihre Ergebnisse hin getestet. Startzeitpunkt war das erstmalige Erscheinen der Meldung im Reuters-Ticker. Die verwendete Suchanfrage lautete:

- Hubschrauber Absturz im Irak

Wertung der Suchergebnisse

Die oben beschriebenen Fragen wurden in den genannten Zeiträumen an die untersuchten Nachrichtensuchmaschinen gestellt. Nur die ersten zehn Suchergebnisse wurden in die Auswertung einbezogen.²⁶ Als Treffer gilt ein Suchergebnis, wenn:

- Titel und Abstract übereinstimmen sowie der Artikel relevant ist;
- entweder Titel oder Abstract nicht übereinstimmen, der Artikel jedoch relevant ist;
- weder Titel noch Abstract im Suchergebnis mit der Meldung übereinstimmen, der Artikel jedoch relevant ist.

Als Nichttreffer wurde gewertet, wenn:

26 Da nach einer Umfrage unter Suchmaschinenbetreibern durchschnittlich nur 4,5 Ergebnis-Links pro Anfrage durch die Nutzer aufgerufen werden (vgl. Machill et al. 2003: 94f.), wurde die Auswertung auf die ersten zehn Suchergebnisse beschränkt. Dies geschah auch vor dem Hintergrund, dass das Spam-Problem bei News-Suchmaschinen ausgeschlossen werden kann, so dass alle zehn Treffer potenziell relevant sein sollten.

- eine Meldung hinsichtlich Anbieter und Inhalt identisch ist mit einer anderen in der Liste des gleichen Durchgangs;
- nicht auf einen Artikel, sondern auf eine Startseite/Portalseite verlinkt wird.
- der Artikel nicht relevant hinsichtlich der Suche ist;
- Titel und Abstract mit der Meldung übereinstimmen, jedoch zu einer nicht relevanten oder nicht existierenden Seite führen.

Als relevant wurden all jene Treffer gewertet, die zur Query, also der verwendeten Suchphrase, unabhängig vom Erscheinungsdatum passende Artikel lieferten. Als relevant und aktuell wurden all jene Treffer gewertet, die den aktuellen Stand der Berichterstattung zeigten und maximal bis zu 24 Stunden alt waren, bzw. den Stand der Nachrichtenlage, der in den Ticker-Meldungen erkennbar war, widerspiegeln.²⁷

3.2 Fünf-Tage-Test

3.2.1. Vorgehen

Zeitlicher Ablauf

Die zehn untersuchten Suchmaschinen wurden über einen Zeitraum von fünf aufeinander folgenden Tagen betrachtet. Nach dem Ereignis startete die erste Suche am darauf folgenden Tag. Die Suchanfrage erfolgte jeweils gegen 11:00 Uhr (Abb.7).

Abb. 7: Untersuchungszeitraum des Fünf-Tage-Test

Mo	Di	Mi	Do	Fr	Sa	So
Ereignis						
	Zeitraum der Untersuchung					

Die Suchergebnisse wurden hinsichtlich Abstract und Titel mit den Daten der Meldung verglichen. Das Datum der Meldung wurde nicht ausgewertet.

Besonderheiten der Suchmaschinen für die Auswertung

Google nutzt eine spezielle Form der Ergebnisdarstellung: Für jeden gefundenen Treffer werden dem ursprünglichen Treffer zusätzlich bis zu drei ähnliche Suchergebnisse zugeordnet. Diese stammen aus weiteren Quellen und fanden im Test aus methodischen Gründen keine Berücksichtigung, da sich auf einer Seite bis zu 40 Ergebnisse finden. Eine Vergleichbarkeit mit den übrigen Suchmaschinen wäre so nicht gegeben. Bei Web.de wurde die News-Suche unter „Schlagzeilen“ benutzt, bei MSN die *Focus*-Suche unter Schlag-

27 Wie in Abschnitt 2.3.2 beschrieben funktioniert die Bestimmung des Nachrichtendatums nicht immer fehlerfrei. Aus diesem Grund wurde der Aspekt der inhaltlichen Aktualitätsbestimmung eingeführt.

zeilen als News-Suchmaschine. Bei Lycos und Fireball wurde die assoziierte Nachrichtensuchmaschine Paperball genutzt.

3.2.2 Die Fragen

Query 1: Schill

Zur Zeit des Tests war der Skandal um den ehemaligen Hamburger Innensenator Ronald Barnabas Schill umfassend in den Medien vertreten. Er hatte dem Bürgermeister Ole von Beust gedroht, er werde Informationen über dessen Homosexualität an die Öffentlichkeit bringen. Wir wollten wissen, inwieweit die aktuelle Entwicklung bei den Suchergebnissen repräsentiert wird. Durch die Vielzahl an Berichten schon über die dem Test vorausgegangenen Monate hinaus war die Gefahr gegeben, dass Suchmaschinen, die das Datum einer Meldung nicht ausreichend berücksichtigen, sehr alte Meldungen als Treffer liefern. Der eindeutige Name barg hingegen keine Verwechslungsgefahr mit anderen Themen.

Query 2: Simbabwe Commonwealth

Anfang Dezember trat Simbabwe aus dem Commonwealth aus. Es war abzu-sehen, dass diese Meldung kaum Folgenachrichten nach sich zieht und insofern schnell aus den Schlagzeilen verschwinden würde. Wir wollten wissen, ob die Nachricht wenigstens eine Woche lang recherchierbar bleibt.

Query 3: Toll Collect Maut

Die Querelen um die LKW-Maut in Deutschland beschäftigten die Medien vor und nach dem Test. Für die Suchmaschinen bestand die Gefahr, bei dieser Frage, die nicht als Phrasensuche formuliert war, Treffer zu Berichten über andere Maut-Themen zu liefern, wenn die Mehrwortsuche nicht korrekt funktioniert hatte. Auch hier ist das Angebot an veralteten Meldungen sehr hoch, ebenso wie die Dynamik der aktuellen Berichterstattung. Wir wollten den aktuellen Stand der Dinge erfahren.

Query 4: Irak Opfer

Die Fragestellung war offen formuliert, so dass sowohl aktuelle Meldungen als auch Dossiers und thematisch und zeitlich weiter gefasste Beiträge als Treffer in Frage kamen. Wir wollten wissen, ob eher aktuelle Meldungen oder stärker verlinkte, aus mehreren Seiten bestehende Texte gefunden werden.

Query 5: Parlamentswahl Russland

Auch dieses Auslandsthema barg viele Verwechslungsmöglichkeiten mit älteren Meldungen und die Gefahr, dass die Mehrwortsuche nicht beide Stichwörter gleichermaßen berücksichtigt. Erwartet wurde eine aktuelle Berichterstattung über den Ausgang der Wahl im Dezember 2003.

3.2.3 Auswertung: Relevante und aktuelle Treffer

Tab. 1.1: Relevante Treffer Schill

	1	2	3	4	5	M
AltaVista	10	10	7	8	5	8
Google	10	8	10	9	10	9,4
Yahoo	5	5	5	5	5	5
Paperball	10	10	7	10	10	9,4
T-Online	0	0	0	0	0	0
Web.de	10	10	10	0	10	8
MSN	3	5	1	3	1	2,6
AOL	0	0	0	0	0	0
Metager	5	5	5	5	5	5

Tab. 1.2: Relevante und aktuelle Treffer Schill

	1	2	3	4	5	M
AltaVista	10	8	6	9	3	7,2
Google	10	7	8	9	8	8,4
Yahoo	3	3	3	3	3	3
Paperball	2	10	6	10	10	7,6
T-Online	0	0	0	0	0	0
Web.de	10	9	3	0	5	5,4
MSN	0	1	1	2	1	1
AOL	0	0	0	0	0	0
Metager	3	3	3	3	3	3

Tab. 1.3: Relevante Treffer Simbabwe Commonwealth

	1	2	3	4	5	M
AltaVista	9	9	9	0	10	7,4
Google	10	9	9	10	10	9,6
Yahoo	10	10	10	10	10	10
Paperball	10	3	0	0	0	2,6
T-Online	0	0	0	0	0	0
Web.de	0	0	0	0	0	0
MSN	1	1	1	0	0	0,6
AOL	6	9	4	5	10	6,8
Metager	2	2	2	2	3	2,2

Tab. 1.4: Relevante und aktuelle Treffer Simbabwe Commonwealth

	1	2	3	4	5	M
AltaVista	7	7	7	0	7	5,6
Google	8	6	6	2	3	5
Yahoo	7	7	7	7	7	7
Paperball	10	3	0	0	0	2,6
T-Online	0	0	0	0	0	0
Web.de	0	0	0	0	0	0
MSN	1	1	0	0	0	0,4
AOL	0	7	0	2	5	2,8
Metager	2	2	2	2	3	2,2

Tab. 1.5: Relevante Treffer Toll Collect Maut

	1	2	3	4	5	M
AltaVista	10	10	9	9	8	9,2
Google	10	10	10	10	9	9,8
Yahoo	8	8	8	8	8	8
Paperball	10	10	10	10	10	10
T-Online	0	4	1	0	1	1,2
Web.de	9	9	9	8	8	8,6
MSN	2	2	2	1	1	1,6
AOL	8	8	8	8	8	8
Metager	3	3	3	2	3	2,8

Tab. 1.6: Relevante und aktuelle Treffer Toll Collect Maut

	1	2	3	4	5	M
AltaVista	10	10	4	5	7	7,2
Google	6	9	10	3	8	7,2
Yahoo	7	7	7	7	7	7
Paperball	7	7	7	7	7	7
T-Online	0	1	0	0	0	0,2
Web.de	3	5	4	0	6	3,6
MSN	2	2	2	1	1	1,6
AOL	0	0	0	0	0	0
Metager	0	0	0	0	0	0

Tab. 1.7: Relevante Treffer Irak Opfer

	1	2	3	4	5	M
AltaVista	9	0	6	8	8	6,2
Google	10	10	10	10	10	10
Yahoo	8	3	3	8	8	6
Paperball	4	3	3	9	9	5,6
T-Online	0	0	0	0	0	0
Web.de	8	8	8	8	8	8
MSN	0	0	0	0	0	0
AOL	4	4	4	4	4	4
Metager	0	0	0	0	1	0,2

Tab. 1.9: Relevante Treffer Parlamentswahlen Russland

	1	2	3	4	5	M
AltaVista	10	3	1	1	8	4,6
Google	10	9	9	9	8	9
Yahoo	8	8	8	8	8	8
Fireball	10	0	0	0	1	2,2
Lycos	10	0	0	0	0	2
T-Online	0	0	0	0	1	0,2
Web.de	9	9	9	8	9	8,8
MSN	4	2	2	1	0	1,8
AOL	9	9	8	8	8	8,4
Metager	2	2	2	2	2	2

Tab. 2.1: Relevante Treffer Gesamt

Gesamt	1	2	3	4	5	M
AltaVista	48	32	32	26	39	35,4
Google	50	46	48	48	47	47,8
Yahoo	39	34	34	39	39	37
Paperball	44	26	20	29	30	29,8
T-Online	0	4	1	0	2	1,4
Web.de	36	36	36	24	35	33,4
MSN	10	10	6	5	2	6,6
AOL	27	30	24	25	30	27,2
Metager	12	12	12	11	14	12,2

Tab. 1.8: Relevante und aktuelle Treffer Irak Opfer

	1	2	3	4	5	M
AltaVista	6	0	0	8	8	4,4
Google	0	0	1	1	0	0,4
Yahoo	8	0	3	2	2	3
Paperball	4	1	1	9	8	4,6
T-Online	0	0	0	0	0	0
Web.de	3	3	3	3	3	3
MSN	0	0	0	0	0	0
AOL	0	0	0	0	0	0
Metager	0	0	0	0	0	0

Tab. 1.10: Relevante und aktuelle Treffer Parlamentswahlen Russland

	1	2	3	4	5	M
AltaVista	10	1	0	0	3	2,8
Google	10	8	4	4	1	5,4
Yahoo	8	8	8	8	8	8
Fireball	10	0	0	0	0	2
Lycos	10	0	0	0	0	2
T-Online	0	0	0	0	0	0
Web.de	9	9	9	0	0	5,4
MSN	4	2	2	1	0	1,8
AOL	0	0	0	0	1	0,2
Metager	0	0	0	0	0	0

Tab. 2.2: Relevante und aktuelle Treffer Gesamt

Gesamt	1	2	3	4	5	M
AltaVista	43	26	17	22	28	27
Google	34	30	29	19	20	26
Yahoo	33	25	28	27	27	28
Paperball	33	21	14	26	25	24
T-Online	0	1	0	0	0	0,2
Web.de	25	26	19	3	14	17
MSN	7	6	5	4	2	4,8
AOL	0	7	0	2	6	3
Metager	5	5	5	5	6	5,2

Abb. 8: Gesamt-Treffer der Fünf-Tage-Suche

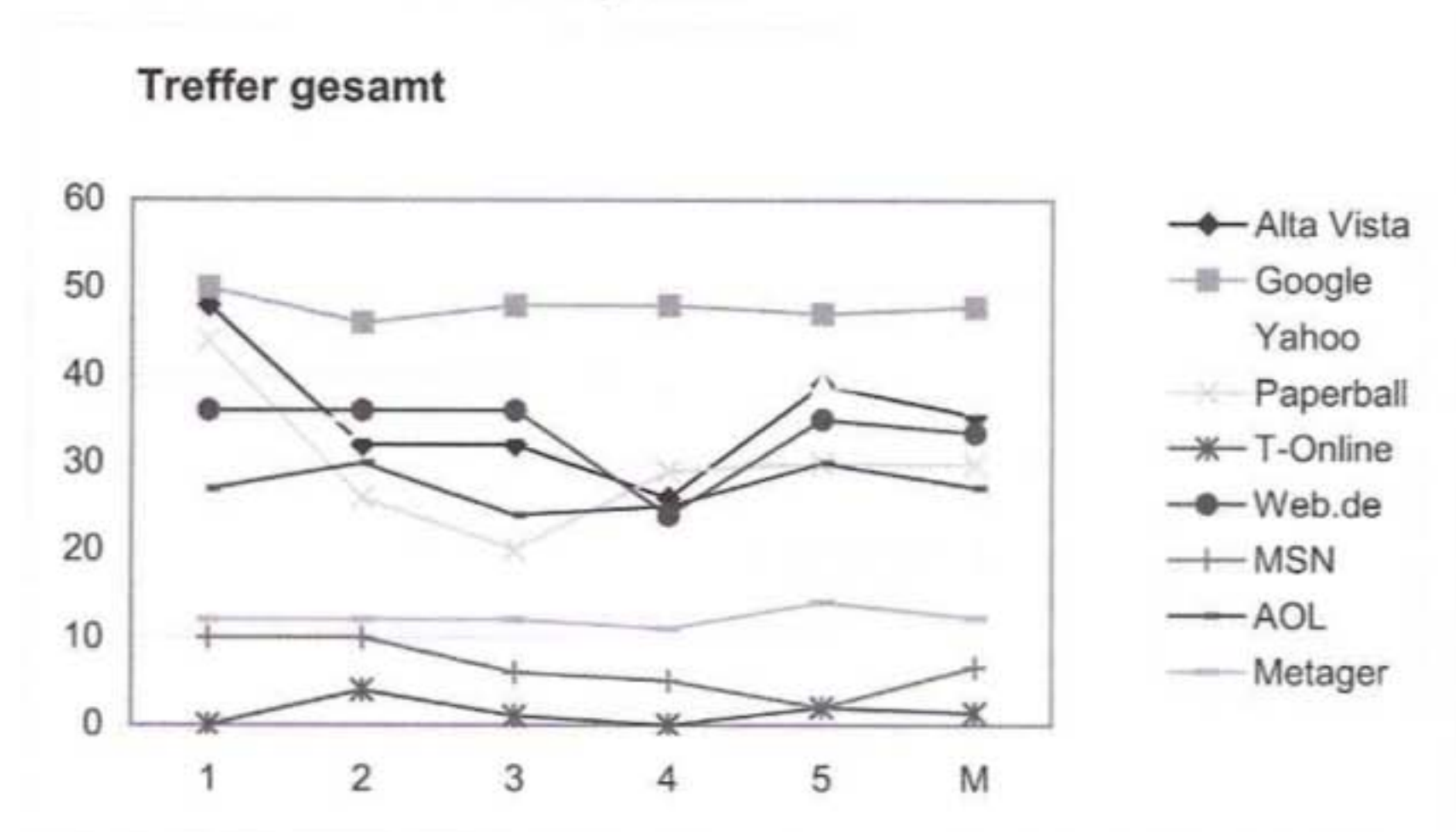
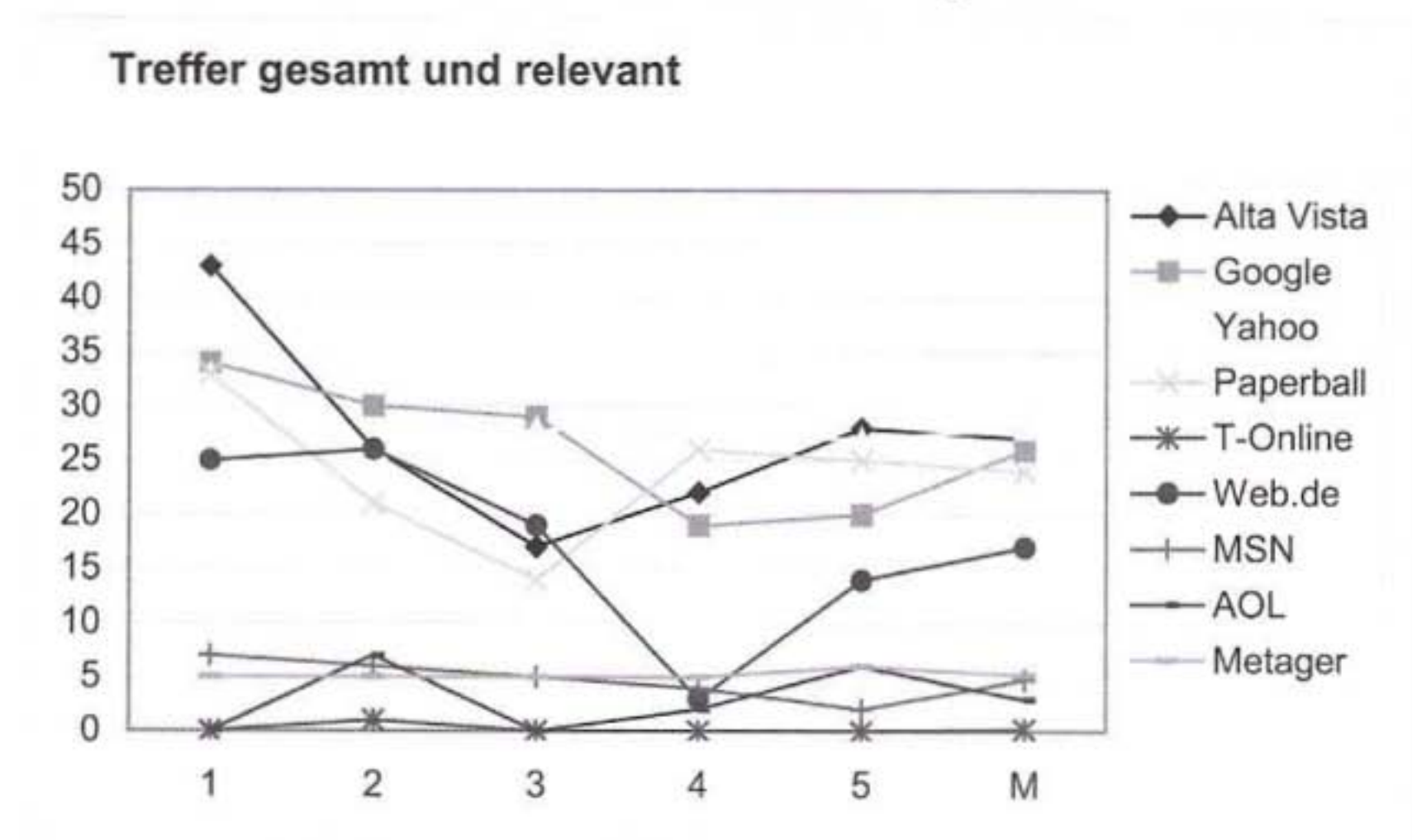


Abb. 9: Relevante und aktuelle Treffer der Gesamt-Fünf-Tage-Suche



Die Gesamtzahl der relevanten Treffer insgesamt bleibt über den Suchzeitraum relativ konstant (Abb. 8). Es ist aber zu beobachten, dass bereits nach einem Tag ein Einbruch bei der Anzahl der Treffer, die sowohl relevant als auch für das Thema aktuell sind, erfolgt (Abb. 9). Dieser Einbruch kann auf zwei mögliche Aspekte zurückgeführt werden. Zum einen ist es möglich, dass es zu den untersuchten Themen keine aktuellen Veröffentlichungen gab. Denkbar ist aber auch, dass die aktuellsten Medien sich in der Rangliste nicht

gegen die älteren Meldungen durchsetzen konnten und somit bei der Betrachtung der ersten zehn Ergebnisse der Eindruck entsteht, es gäbe nichts Neues mehr.

AltaVista-News, Google News und Yahoo bieten die höchste Trefferquote und Aktualität. Die T-Online- News-Suche ist dagegen nahezu unbrauchbar. Da AOL den Google-Index nutzt, gibt es hier zwar gute Treffer zu Themen, die schon eine Weile publiziert und daher bereits im Google-Index zu finden sind. Durch den nur monatlich stattfindenden Aktualisierungszyklus sind aktuelle Meldungen jedoch nicht zu finden. MSN verweist ausschließlich auf Partnerangebote wie Focus und findet nur eine sehr geringe Anzahl an Treffern. Titel und Abstract sind häufig nichtssagend. Auch werden Inhaltsverzeichnisse alter *Focus*-Ausgaben und die *Focus*-Startseite als Treffer angeboten.

Auswertung der einzelnen Suchanfragen

Query 1: Schill

Die Treffer zum Skandal um den ehemaligen Hamburger Innensenator spiegeln die aktuelle Berichterstattung. Nur wenige Treffer verwiesen auf allgemeine Seiten, die keine Informationen zu der aktuellen Entwicklung enthielten.

Beispiele für gute Treffer:

- Schill-Partei: Ole von Beust will Hamburger Bürgerschaft auflösen
- Hamburger Koalition geplatzt
- Schill-Partei will Schill ausschließen
- Nach Bruch der Hamburger Regierungskoalition: Bundesvorstand entmachtet Schill

Beispiele für unpassende Treffer:

- dpa-Nachrichtenüberblick Politik
- Europa winkt ab
- Harald geht
- Terror, Medien und 68er

Query 2: Simbabwe Commonwealth

Da Simbabwe relativ selten in der aktuellen Berichterstattung auftaucht, waren die Treffer, die nicht auf das aktuelle Ereignis verwiesen solche, die allgemeine Informationen über das Land enthielten. Die Treffer befassten sich erwartungsgemäß mit dem Austritt Simbawes und dem Regierungschef.

Beispiele für aktuelle Treffer:

- Mugabe in der Isolation
- tagesschau.de: Commonwealth-Gipfel streitet über Simbabwe

- Simbabwe aus Commonwealth ausgetreten
- Austritt aus Commonwealth: Simbabwe weiter auf Isolationskurs - ...

Inaktuelle Treffer:

- Internationale Strafe | Afrika | Deutsche Welle | 20.03.2002
- Enteignung und Umverteilung von Land in Simbabwe, 04.09.2002
- Bericht aus Simbabwe, Konrad-Adenauer-Stiftung e.V.
- Afrika, Diamanten

Query 3: Toll Collect Maut

Die Ergebnisse behandelten zum großen Teil den aktuellen Entwicklungsstand. Etliche Ergebnisse verwiesen auf allgemeine Artikel oder solche, die einen früheren Stand der Entwicklung dokumentieren.

Beispiele für aktuelle Treffer:

- Stolpe verlangt mehr Geld von Toll Collect
- Maut-Desaster Stolpe will mehr Geld von Toll Collect
- Grüne: Bund muss Maut-Schadenersatz notfalls gerichtlich erzwingen
- Toll Collect: Fahrt in die Irre – manager-magazin.de

Inaktuelle Treffer:

- Börsentag auf einen Blick: Börsianer rechnen mit schwächerem DAX-Start
- Verhandlungsrunde des Ministeriums mit Toll Collect gescheitert
- Telekom will Probleme bei Lkw-Maut „in Ordnung bringen“
- Ministerium – Kein Maut-Gipfel diese Woche vereinbart

Query 4: Irak Opfer

Da es im Irak täglich Opfer zu verzeichnen gibt, werden auch eine Vielzahl entsprechender Meldungen publiziert. Die nicht aktuellen Treffer haben erwartungsgemäß ältere Meldungen und Hintergrundberichte zur Entwicklung im Irak zum Thema.

Beispiele für aktuelle Treffer:

- Ersatz für ermordetes irakisches Ratsmitglied ernannt
- US-Soldat und irakischer Polizist getötet
- Stille Killer
- Explosionen im Osten Bagdads

Inaktuelle Treffer:

- Annan ernennt Mountain zum UN-Gesandten für Irak
- Ministerpräsident Koizumi ruft japanische Bevölkerung zur Unterstützung der Wiederaufbaubemühungen im Irak und Entsendung von japanischen Streitkräfte auf
- Rock & Pop Splitter

- Brot für die Welt Hamburg

Query 5: Parlamentswahl Russland

Auch hier wurden bei den aktuellen Treffern die Erwartungen erfüllt. Die inaktuellen oder Fehltreffer waren teilweise Startseiten von Portalen und veraltete Nachrichtenübersichten.

Beispiele für aktuelle Treffer:

- US-Regierung besorgt über Wahl in Russland
- Aktien Moskau Schluss: Gut behauptet nach Parlamentswahlen
- Kreml-treue Parteien gewinnen Dumawahl
- Zweidrittel-Mehrheit für Putin-Parteien

Inaktuelle Treffer:

- tagesschau.de: 05.12.2003
- Abacho.de – Startseite
- Polen (EUROPA Info online)
- Metro-Chef Körber: Chancen des russischen Marktes nicht verschlafen

3.3 24-Stunden-Test: Vergleich News-/Websucher +10h und +24h

3.3.1 Vorgehen

Zeitlicher Ablauf

Die zehn untersuchten Suchmaschinen wurden 10 Stunden und 24 Stunden nach einem Ereignis parallel getestet. Als Ereigniszeitpunkt wurden Datum und Uhrzeit der ersten Meldung im Reuters-Ticker festgehalten. Die normale Websuche und die spezielle Nachrichtensuche wurden gleichermaßen betrachtet. Dies spiegelt sich in den Ergebnissen „Treffer“ und „Treffer News“ wider (vgl. Tab.3). Die Fragen konnten somit nicht wie in den anderen Testblöcken zeitgleich gestellt werden. Vielmehr musste jeweils ein passendes Ereignis abgewartet werden.

Besonderheiten der Suchmaschinen für die Auswertung

Metager und AOL verfügen nicht über eine eigene News-Suche. Daher konnte hier kein Vergleich stattfinden. Die Fireball-Nachrichtensuche und die Lycos-Nachrichtensuche wurden über Paperball abgewickelt und sind daher identisch.

Die Fragen

Query 1: Explosion Bagdad

Aktuelles Ereignis war folgende Reuters-Meldung:

Polizei: 17 Tote bei Bombenexplosion in Bagdad

Bagdad (Reuters) – Bei einer Bombenexplosion in der irakischen Hauptstadt Bagdad sind nach Polizeiangaben am Mittwoch mindestens 17 Menschen getötet worden.

Query 2: Schill

Aktuelles Ereignis war folgende Reuters-Meldung:

Hamburger Ex-Innensenator Schill aus Partei ausgeschlossen

Der Parteigründer Ronald Schill ist aus der Partei Rechtsstaatlicher Offensive – auch bekannt als Schill-Partei – ausgeschlossen worden.

Query 3: Flugpassagierdaten

Aktuelles Ereignis war folgende Reuters-Meldung:

EU zu Austausch von Flugpassagierdaten mit den USA bereit

Im Streit um den Austausch von Luftverkehrsdaten haben die USA und die EU-Kommission eine Einigung erzielt. Damit soll der von den USA angeführte Kampf gegen den Terror unterstützt werden

Query 4: Vermittlungsausschuss

Aktuelles Ereignis war folgende Reuters-Meldung:

Vermittlungsausschuss verabschiedet Reform-Kompromiss

Berlin (Reuters) – Der Vermittlungsausschuss von Bundestag und Bundesrat hat am Dienstag kurz vor Mitternacht dem Kompromiss zum Vorziehen von Steuerentlastungen und weiteren Reformen im Steuer- und Arbeitsmarktbereich beschlossen.

Query 5: SARS Taiwan

Aktuelles Ereignis war folgende Reuters-Meldung:

Erstmals wieder Sars-Fall in Taiwan

Taipeh (Reuters) – Erstmals seit Monaten gibt es in Taiwan wieder einen Fall der Lungenkrankheit Sars. Börsen in Asien reagierten auf die Nachricht mit Kursverlusten.

Als Antworten wurden Beiträge erwartet, die die in den Meldungen dargestellten Sachverhalte thematisierten, keine älteren Treffer mit ähnlichen Themen. Nur relevante und aktuelle Ergebnisse wurden als Treffer gewertet. Die Auswahl soll wie bereits beim Fünf-Tage-Test (vgl. Abschnitt 3.2.2) einen Querschnitt der möglichen Probleme abdecken: Verwechslungsgefahr mit älteren Nachrichten zum gleichen Thema, schnelle Fortentwicklung des Themas, Funktion der Mehrwortsuche, regionale, nationale und internationale Fragestellung.

3.3.2 Auswertung

Zahl der Treffer insgesamt

Die Tabelle listet die Treffer zu allen fünf durchgeführten Queries, im ersten Teil aufgeteilt nach normaler Suche und Suche über die Nachrichtensuche. Den Vergleich nach Zeitpunkt der Suche bietet der zweite Tabellenteil. Die maximale Trefferzahl beträgt bei fünf Fragen und zehn angezeigten Treffern somit 50 pro Suchmaschine.

Tab. 3: Treffer 24h-Suche nach Zeit und nach Art der Suche

Treffer	Vergleich nach Zeit									
	Alta Vista	Fireball	Lycos	Google	Web.de	Yahoo	AOL	MSN	T-Online	Metager
Treffer +10	3	8	2	0	2	2	2	3	2	0
Treffer +24	1	13	11	10	9	8	2	6	2	0
Treffer News-Suche +10	28	35	35	39	29	28	0	5	0	0
Treffer News-Suche +24	32	34	34	41	28	30	0	5	0	0

Treffer	Vergleich nach Art der Suche									
	Alta Vista	Fireball	Lycos	Google	Web.de	Yahoo	AOL	MSN	T-Online	Metager
Treffer +10	3	8	2	0	2	2	2	3	2	0
Treffer News-Suche +10	28	35	35	39	29	28	0	5	0	0
Treffer +24	1	13	11	10	9	8	2	6	2	0
Treffer News-Suche +24	32	34	34	41	28	30	0	5	0	0

Abb. 10.1: Treffer 24h-Suche im zeitlichen Verlauf

Treffer (normale Suche)

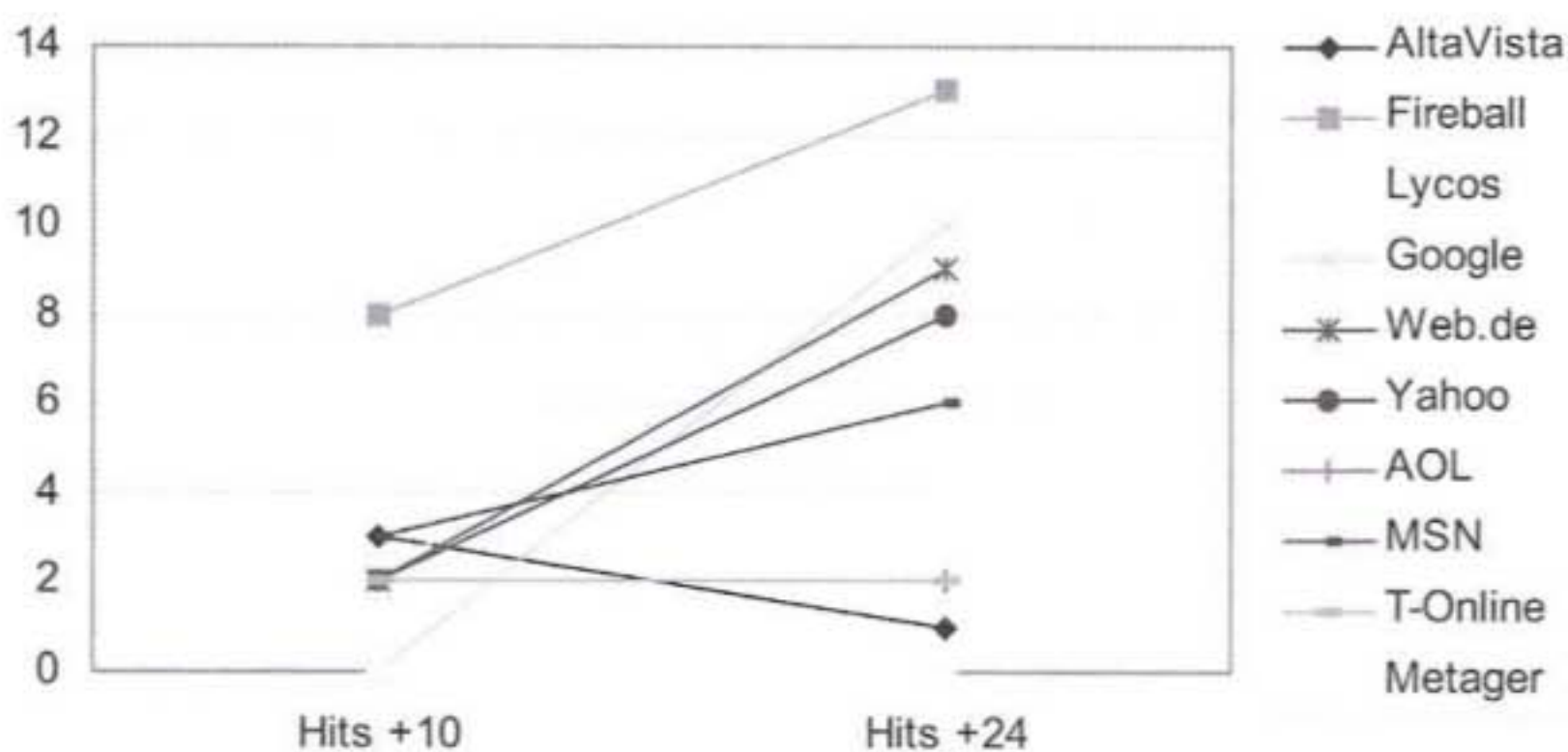
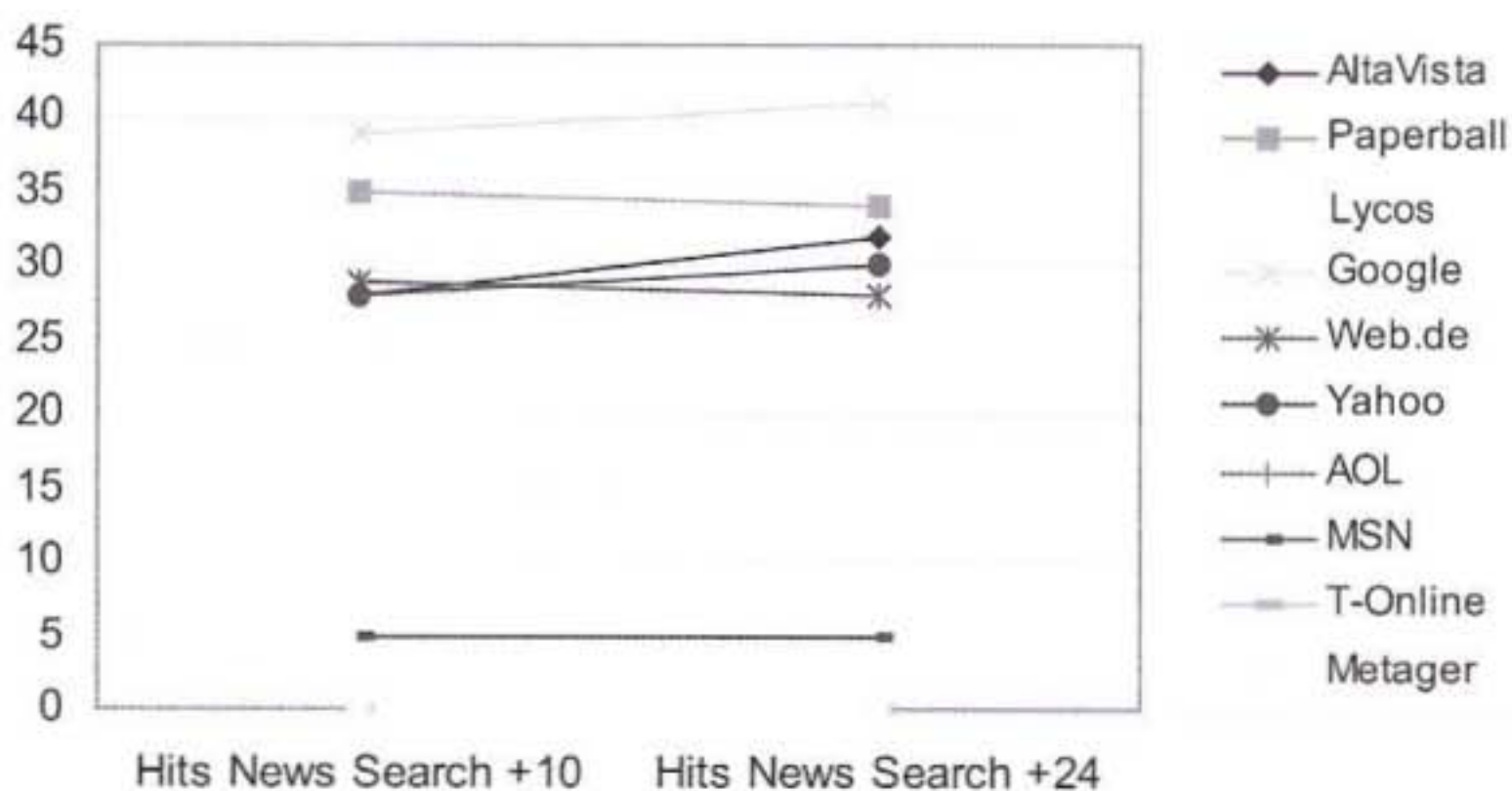


Abb. 10.2: Treffer 24h-Nachrichtensuche im zeitlichen Verlauf

Treffer (Nachrichtensuche)



Erwartungsgemäß führte die Suche in den Nachrichtensuchmaschinen zu deutlich höheren Trefferquoten. Mit einer Verzögerung von 24 Stunden wurde auch die Trefferrate der normalen Suche deutlich besser (Abb. 10.1). Auch die Nachrichtensuche liefert nach diesem Zeitabschnitt die besseren Ergebnisse (Abb. 10.2). Die Fireball-Websuche findet besonders schnell eine Viel-

zahl von guten Treffern, gefolgt von AltaVista und MSN. Die Nachrichtensuchen von Google und Paperball liefern nach zehn Stunden die besten Ergebnisse, ihnen folgen Web.de, AltaVista und Yahoo. T-Online, Metager und MSN versagen vollständig. Die T-Online-Nachrichtensuche ist teilweise schlechter als die normale Websuche des Portals.

Tab. 4: Übereinstimmung im Titel – 24h-Suche

Titel	Vergleich nach Zeit									
	Alta Vista	Fireball	Lycos	Google	Web.de	Yahoo	AOL	MSN	T-Online	Metager
Titel +10	1	7	0	2	2	2	2	1	2	0
Titel +24	1	4	0	2	2	2	2	2	2	20
Titel News +10	18	21	19	24	21	18	0	0	0	0
Titel News +24	15	21	19	21	21	18	0	0	0	0

Titel	Vergleich nach Art der Suche									
	Alta-Vista	Fireball	Lycos	Google	Web.de	Yahoo	AOL	MSN	T-Online	Metager
Titel +10	1	7	0	2	2	2	2	1	2	20
Titel News +10	18	21	19	24	21	18	0	0	0	0
Titel +24	1	4	0	2	2	2	2	2	2	2
Titel News +24	15	21	19	21	21	18	0	0	0	0

Tab. 5: Übereinstimmung im Abstract – 24h-Suche

Abstract	Vergleich nach Zeit									
	Alta Vista	Fireball	Lycos	Google	Web.de	Yahoo	AOL	MSN	T-Online	Metager
Abstract +10	2	7	1	2	2	2	2	0	2	0
Abstract +24	4	4	1	2	2	2	2	1	2	20
Abstract News-Suche+10	28	34	34	37	0	25	0	2	0	0
Abstract News +24	24	33	33	37	0	23	0	2	0	0

Abstract	Vergleich nach Art der Suche									
	Alta Vista	Fireball	Lycos	Google	Web.de	Yahoo	AOL	MSN	T-Online	Metager
Abstract	2	7	1	2	2	2	2	0	2	20
Abstract News- Suche	28	34	34	37	0	25	0	2	0	0
Abstract +24	4	4	1	2	2	2	2	1	2	0
Abstract News +24	24	33	33	37	0	23	0	2	0	0

Die Übereinstimmungen in Titel (Tab. 4) und Abstract (Tab. 5) liegen deutlich unter der Zahl der Gesamttreffer (Tab. 3). Abbildung 10 verdeutlicht diesen Zusammenhang. Der Grund sind die teilweise häufigen Überarbeitungen der einzelnen Meldungen. Bei gleich bleibender URL werden entweder die Meldungen der Agenturen auf den aktuellen Stand gebracht oder vom Nachrichtenanbieter selbst verfasste Artikel um neue Erkenntnisse ergänzt. Der Abstract stimmt häufiger überein als der Titel, da hier oft der Title-Tag der HTML-Seite und nicht der Titel der Meldung genutzt wird. Ein weiterer Grund sind Meldungen, die mit einem „Universaltitle“ überschrieben sind, etwa „Reisen, Wetter, Weltweit, MSN, Focus-Online, Asien, Irak, Mosul“.

3.4 Reaktionstest: Vergleich News-/Websucher +1h +3h +5h

3.4.1 Vorgehen

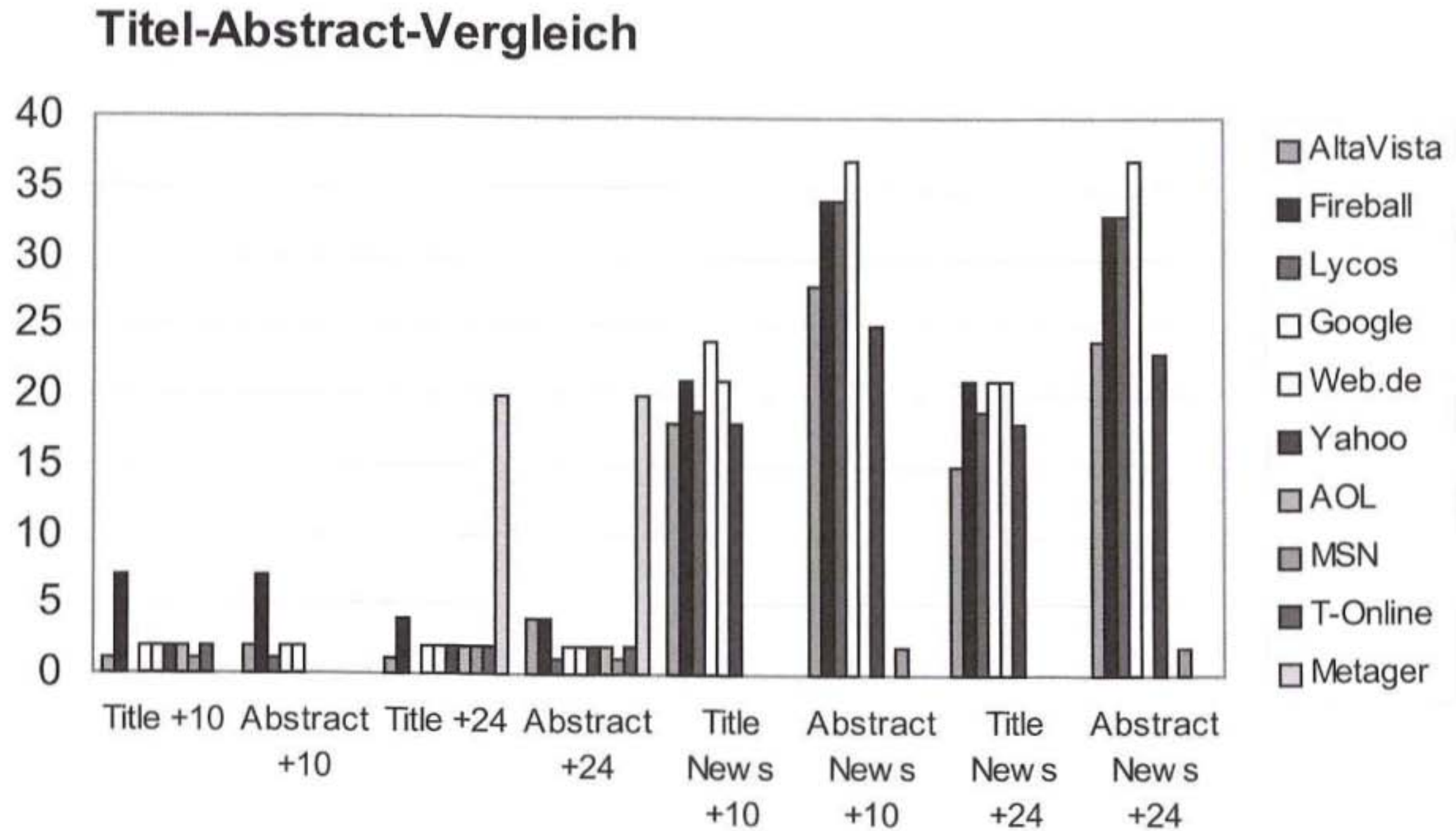
Zeitlicher Ablauf

Neun Suchmaschinen wurden eine, drei und fünf Stunden nach einem von Reuters gemeldeten Zeitpunkt eines Ereignisses auf ihre Ergebnisse hin getestet. Startzeitpunkt war das erstmalige Erscheinen der Meldung im Reuters-Ticker.

Als Suchanfrage wurde das aktuelle Ereignis eines Hubschrauberabsturzes im Irak gewählt. Die Auswahl erfolgte nach ähnlichen Kriterien wie in 3.3.1. Da bei einem solchen Ereignis in den ersten Stunden nach Eintreffen stetig neue Erkenntnisse und Spekulationen über Hergang und mögliche Ursachen über die Ticker verbreitet werden, kann anhand dieser Suchanfrage die Aktualität in Stunden relativ gut verfolgt werden. Auch hier ist die Gefahr der Verwechslung mit älteren Meldungen gegeben.

Die Meldung zum Hubschrauberabsturz im Irak wurde erstmals am Donnerstag, den 8. Januar 2004, 13:49 Uhr CET (Central European Time) veröffentlicht. Die Fragen wurden dementsprechend um 14:49 Uhr, 16:49 Uhr und 18:49 Uhr CET gestellt.

Abb. 11: Vergleich von Titel und Abstract innerhalb der 24h-Suche



Besonderheiten der Suchmaschinen für die Auswertung

Metager und AOL verfügen über keine eigene Nachrichtensuche und wurden daher ausgeschlossen. Zudem wurden von Metager nach zehn Stunden Reaktionszeit schon beim 24-Stunden-Test (vgl. Abschnitt 3.3) keine Treffer gefunden. Zusätzlich wurde die unabhängige News-Suchmaschine Paperazzi herangezogen.

3.4.2 Auswertung

Die Tabellen 6.1, 6.2 und 6.3 listen die Treffer zur durchgeführten Suchanfrage unterteilt nach Gesamttreffern, Treffern nach Titel und Treffern nach Abstract auf. In die Auswertung wurden jeweils nur die ersten zehn Suchergebnisse einbezogen.

Tab. 6.1: Treffer Gesamt – Reaktionstest

Titel	Treffer +1h	Treffer +3h	Treffer +5h
MSN	0	1	1
T-Online	0	0	0
Web.de	0	3	5
Yahoo	3	6	6
Paperball (Lycos/Fireball)	0	4	4
Paperazzi	1	2	3

Tab. 6.2: Treffer im Titel – Reaktionstest

Titel	Treffer +1h	Treffer +3h	Treffer +5h
AltaVista	4	8	8
Google	8	8	8
MSN	0	0	0
T-Online	0	0	0
Web.de	0	3	5
Yahoo	3	6	5
Paperball (Lycos/Fireball)	1	5	5
Paperazzi	1	2	3

Tab. 6.3: Treffer im Abstract – Reaktionstest

Abstracts	Treffer +1h	Treffer +3h	Treffer +5h
AltaVista	7	10	8
Google	6	6	6
MSN	0	1	2
T-Online	0	0	0
Web.de	0	0	0
Yahoo	3	6	5
Paperball (Lycos/Fireball)	1	5	5
Paperazzi	1	2	3

Anmerkungen

Sowohl Google, AltaVista und Yahoo liefern mit einer Reaktionszeit von weniger als einer Stunde die besten Ergebnisse. Alle getesteten Suchmaschinen, mit Ausnahme von T-Online, konnten drei Stunden nach der eingegangenen Meldung erste Ergebnisse liefern. MSN lieferte nur ein Ergebnis, dieses war jedoch relevant und verwies auf einen *Focus*-Artikel. Obwohl dieser Artikel bereits nach einer Stunde von anderen Nachrichtensuchern gefunden wurde, war die MSN-/Focus-hauseigene Suchmaschine dazu nicht in der Lage. Der Artikel war erst nach drei Stunden über MSN auffindbar.

3.5 Untersuchung der Ranking-Kriterien bei der Nachrichtensuche: Verbreitungsgebiet, Verflechtungen

3.5.1. Analyse der angezeigten Anbieter

Ziel der Betrachtung war es, herauszufinden, ob es Häufungen der Trefferzahl bei bestimmten Anbietern gibt. Als Grundlage dienen die Daten aus dem Fünf-Tage-Test. Hierbei wurden jedoch nicht alle Ergebnisse berücksichtigt: Yahoo verlinkte ausschließlich auf sein eigenes News-Portal. Web.de verwendete ausschließlich dpa-Meldungen. MSN verwies ebenso ausschließlich auf Meldungen aus dem MSN-Netzwerk. T-Online erbrachte insgesamt nur sieben Treffer und fiel daher ebenfalls für eine Betrachtung aus. AOL und Metager verfügten lediglich über eine normale Websuche und lieferten geringe Treffer.

Somit eignen sich für eine Betrachtung der Streuung der Nachrichtenquellen lediglich die Ergebnisse von AltaVista, Google, und Paperball (für Lycos und Fireball). Die Trefferlisten für diese Anbieter – sortiert nach Quellen – sind in Tabelle 7 aufgeführt.

Tab. 7: Streuung der Treffer bei AltaVista, Google und Paperball

Tref-ferzahl	AltaVista	%	Tref-ferzahl	Google	%	Tref-ferzahl	Paperball	%
23	de.news.yahoo.com	13,1	15	www.abendblatt.de	6,2	20	www.taz.de	15,0
23	www.netzeitung.de	13,1	14	www.dw-world.de	5,9	16	www.vwd.de	12,0
19	www.sueddeutsche.de	10,9	11	www.aller-zeitung.de	4,6	12	www.spiegel.de	9,0
17	www.welt.de	9,7	9	morgenpost.berlin1.de	3,7	10	www.jungewelt.de	7,5
10	www.tagesschau.de	5,7	8	www.bkz-online.de	3,3	10	www.netzeitung.de	7,5
10	www.wams.de	5,7	7	www.az.com.na	2,9	7	www.fr-aktuell.de	5,2
9	www.faz.net	5,1	6	de.news.yahoo.com	2,5	6	www.ftd.de	4,5
8	focus.msn.de	4,6	6	ebund.ch	2,5	6	www.stern.de	4,5
6	www.tagesspiegel.de	3,4	6	russslandonline.ru	2,5	6	www.welt.de	4,5
5	www.baz.ch	2,8	6	www.baz.ch	2,5	5	www.faz.net	3,7
5	www.n-tv.de	2,8	6	www.ebund.ch	2,5	5	www.sueddeutsche.de	3,7
4	www.heise.de	2,3	5	news.tirol.com	2,0	5	www.tagesschau.de	3,7
4	www.heute.T-Online.de	2,3	5	www.bielertagblatt.ch	2,0	3	morgenpost.berlin1.de	2,5
3	www.bayreuth.de	1,7	5	www.fr-aktuell.de	2,0	3	www.abendblatt.de	2,5
3	www.ftd.de	1,7	5	www.nachrichten.at	2,0	3	www.nzz.ch	2,2
2	news.onvista.de	1,1	5	www.netecho.info	2,0	2	www.bund.ch	1,5
2	www.computerwoche.de	1,1	5	www.rundschau-online.de	2,0	2	www.heise.de	1,5
2	www.hrw.org	1,1	5	www.tagesschau.de	2,0	2	www.kleinezeitung.at	1,5
2	www.merkur-online.de	1,1	4	de.internet.com	1,7	2	www.nd-online.de	1,5
2	www.nd-online.de	1,1	4	www.bremer-nachrichten.de	1,7	2	www.zeit.de	1,5
2	www.nzz.ch	1,1	4	www.nzz.ch	1,7	2	www1.news.ch	1,5
2	www.rtl.de	1,1	4	www.sfdrs.ch	1,7	1	news.focus.msn.de	0,8
2	www.1.news.ch	1,1	4	www.solms-braunfelser.de	1,7	1	www.berlinonline.de	0,8
2	www.berlin.sat1.de	1,1	4	www.sueddeutsche.de	1,7	1	www.ostsee-zeitung.de	0,8
1	finanzen.sueddeutsche.de	0,6	3	derstandard.at	1,2	1	www.rp-online.de	0,8
1	www.manager-magazin.de	0,6	3	newswelt.stimme.de	1,2	1	www.wienerzeitung.at	0,8
1	www.nachrichten.at	0,6	3	www.jungewelt.de	1,2			
1	www.orf.at	0,6	3	www.merkur-online.de	1,2			
1	www.rp-online.de	0,6	3	www.ndrinfo.de	1,2			
1	www.stern.de	0,6	3	www.spiegel.de	1,2			
1	www.vol.at	0,6	3	www.weilburger-tageblatt.de	1,2			
1	www.2.onnachrichten.T-Online.de	0,6	3	www.welt.de	1,2			
			2	news.search.ch	0,9			
			2	onnachrichten.T-Online.de	0,9			
			2	www.computerwoche.de	0,9			

Tref- ferzahl	AltaVista	%	Tref- ferzahl	Google	%	Tref- ferzahl	Paperball	%
			2	www.dradio.de	0,9			
			2	www.faz.net	0,9			
			2	www.kurier.at	0,9			
			2	www.mittelhessen.de	0,9			
			2	www.n24.de	0,9			
			2	www.nd-online.de	0,9			
			2	www.netzeitung.de	0,9			
			2	www.oltnertagblatt.ch	0,9			
			2	www.rheinpfalz.de	0,9			
			2	www.swissinfo.org	0,9			
			2	www.berlin.sat1.de	0,9			
			1	finan- zen.sueddeutsche.de	0,4			
			1	information.bluewin.ch	0,4			
			1	onnachrichten.T- Online.de	0,4			
			1	web.saint-paul.lu	0,4			
			1	www.autobild.de	0,4			
			1	www.autosieger.de	0,4			
			1	www.bbv-net.de	0,4			
			1	www.berlinonline.de	0,4			
			1	www.club-transport.de	0,4			
			1	www.die-tagespost.de	0,4			
			1	www.ftd.de	0,4			
			1	www.general-anzeiger- bonn.de	0,4			
			1	www.heise.de	0,4			
			1	www.ktz.at	0,4			
			1	www.nd-online.de	0,4			
			1	www.ndr.de	0,4			
			1	www.news.de	0,4			
			1	www.ngo-online.de	0,4			
			1	www.nn-online.de	0,4			
			1	www.nw-news.de	0,4			
			1	www.orb.de	0,4			
			1	www.orf.at	0,4			
			1	www.paz-online.de	0,4			
			1	www.presseportal.de	0,4			
			1	www.rp-online.de	0,4			
			1	www.rtl.de	0,4			

Tref-ferzahl	AltaVista	%	Tref-ferzahl	Google	%	Tref-ferzahl	Paperball	%
			1	www.seite.com	0,4			
			1	www.stern.de	0,4			
			1	www.suedostschweiz.ch	0,4			
			1	www.T-Online.at	0,4			
			1	www.taz.de	0,4			
			1	www.tvmatrix.de	0,4			
			1	www.wiesbadener-kurier.de	0,4			
			1	www.zol.ch	0,4			
			1	www2.onnachrichten.T-Online.de	0,4			
			1	www3.hna.de	0,4			

Es zeigt sich eine sehr breite Streuung der Treffer-Hosts. AltaVista fand mit jeweils 13 Prozent der Gesamttreffer besonders die Meldungen der Angebote von Netzeitung und Yahoo. Ein Viertel aller Suchergebnisse stammt aus lediglich zwei Quellen. Paperball fand die meisten Treffer in der *taz* und in den Vereinigten Wirtschaftsdiensten. Auch hier stammen 25 Prozent aller Treffer aus zwei Nachrichtenquellen. Google beweist bei der Nachrichtensuche mit insgesamt 83 verschiedenen Quellen die größte Vielfalt. Die sieben meistgenutzten Quellen machen dabei einen Anteil von rund 25 Prozent aus.

Die geringe Quellenvielfalt von AltaVista und Paperball zeigt sich auch, wenn man die Top zehn der Suchergebnislieferanten betrachtet. Diese machen bei AltaVista 78 Prozent der Gesamttreffer aus (Abb. 12), bei Paperball sind es immerhin noch 76 Prozent (Abb. 14). Google beweist an dieser Stelle eine hohe Vielfalt: Nur 38 Prozent der Gesamttreffer stammen von den Top-Ten-Quellen (Abb. 13).

Tab. 8: Top-Ten- Quellen AltaVista

de.news.yahoo.com	13%
www.netzeitung.de	13%
www.sueddeutsche.de	11%
www.welt.de	10%
www.tagesschau.de	6%
www.wams.de	6%
www.faz.de	5%
focus.msn.de	5%
www.tagesspiegel.de	3%
www.baz.ch	3%
remaining providers	25%
Gesamt	100%

Tab. 9: Top-Ten-Quellen Google

www.abendblatt.de	5%
www.dw-world.de	6%
www.aller-zeitung.de	5%
morgenpost.berlin.de	4%
www.bkz-online.de	3%
www.az.com.na	3%
de.news.yahoo.com	3%
ebund.ch	3%
ruslandonline.ru	4%
www.baz.ch	3%
remaining providers	62%
Gesamt	100%

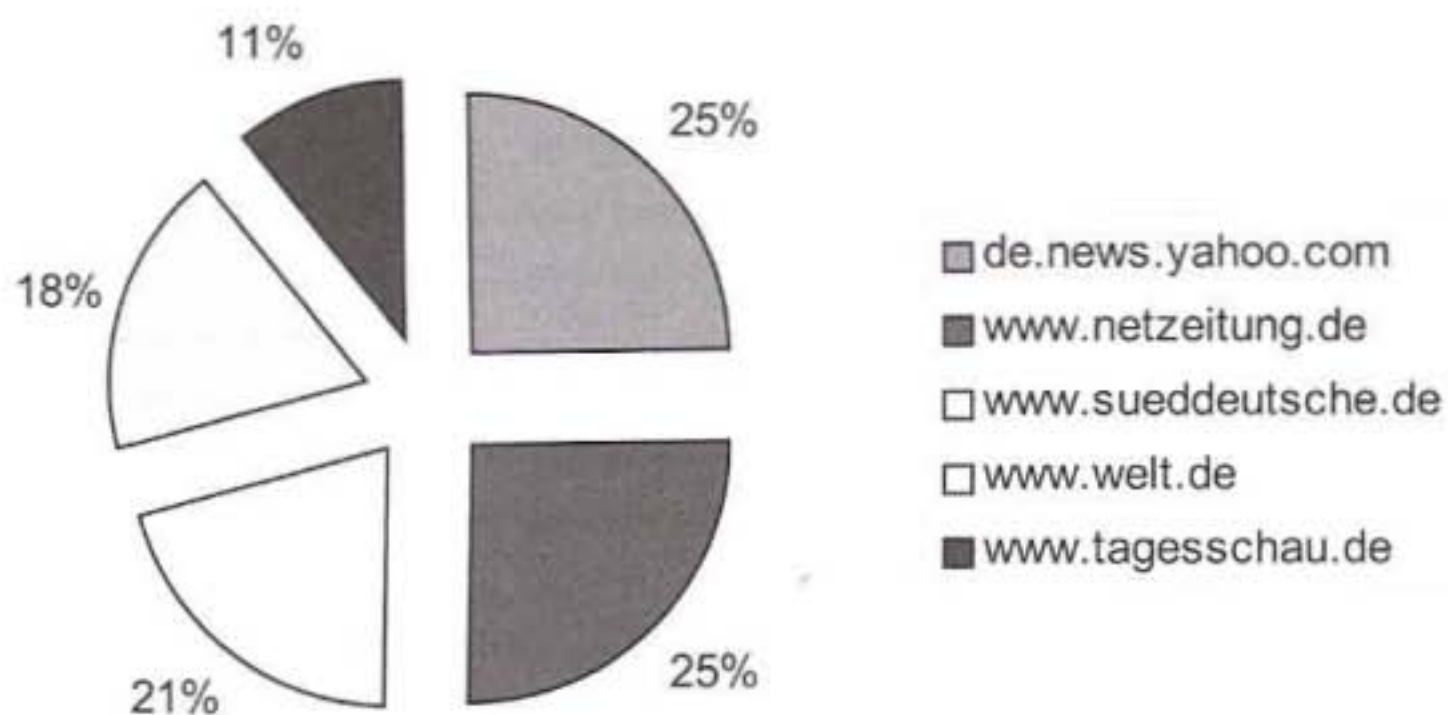
Tab. 10: Top-Ten-Quellen Paperball

www.taz.de	14%
www.vwd.de	11%
www.spiegel.de	9%
www.jungewelt.de	8%
www.netzeitung.de	8%
www.fr-aktuell.de	5%
www.ftd.de	5%
www.stern.de	5%
www.welt.de	5%
www.faz.de	4%
remaining providers	25%
Gesamt	100%

Noch deutlicher wird die Quellenkonzentration, wenn man von der Gesamttrefferzahl die oberen 50 Prozent, also jene Quellen mit den meisten Treffern, betrachtet. AltaVista und Paperball beziehen bei dieser Betrachtungsweise über die Hälfte ihrer Suchergebnisse von lediglich je fünf Quellen (vgl. Abb. 15 und 17). Google streut die Ergebnisse über immerhin 16 verschiedene Anbieter (vgl. Abb. 16).

Abb. 12: 50 Prozent Cut Off AltaVista

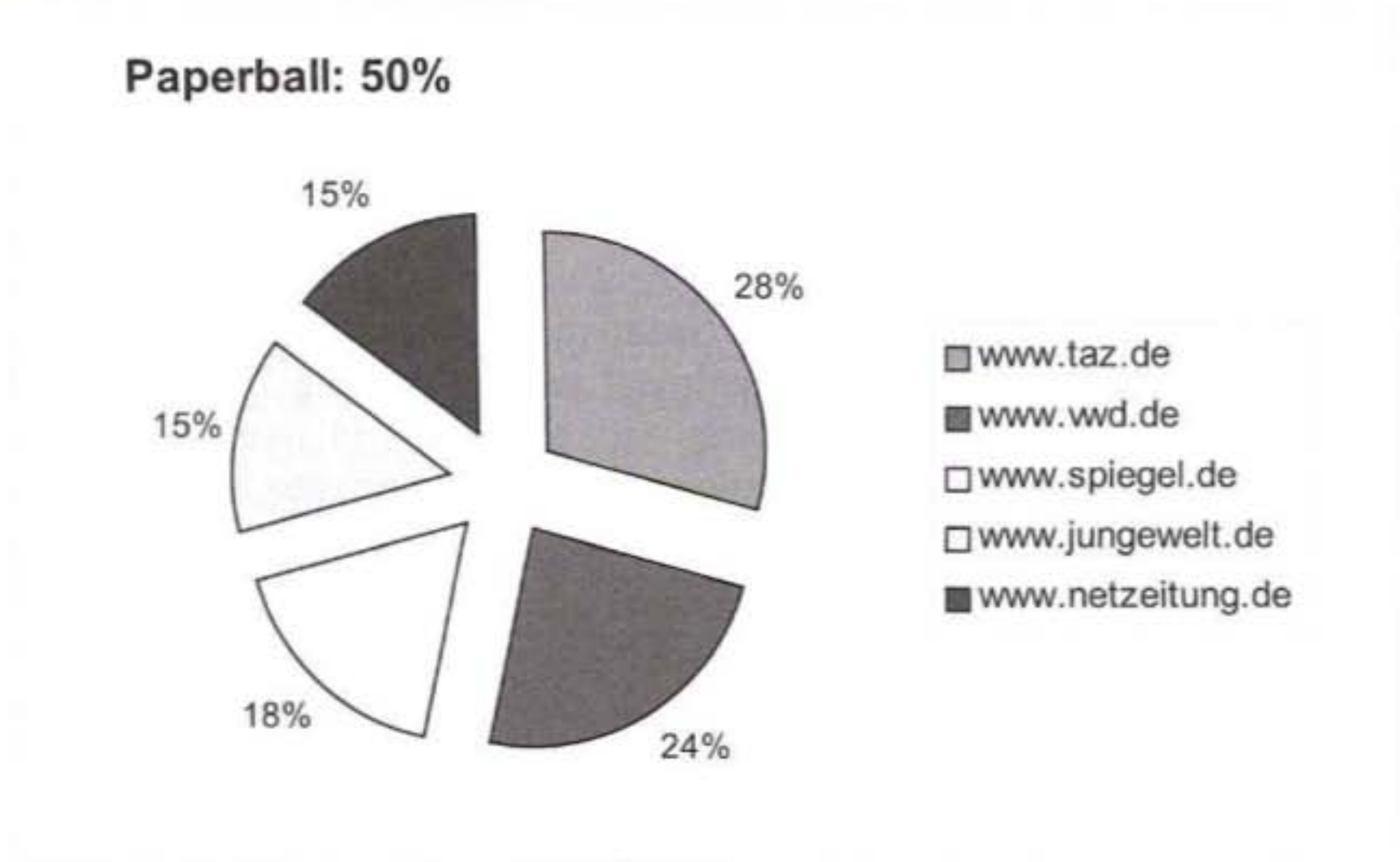
Alta Vista: 50%



Tab. 11: 50 Prozent Cut Off Google

www.abendblatt.de	13%
www.dw-world.de	12%
www.aller-zeitung.de	9%
morgenpost.berlin.de	8%
www.bkz-online.de	7%
www.az.com.na	6%
de.news.yahoo.com	5%
ebund.ch	5%
russslandonline.ru	5%
www.baz.ch	5%
www.ebund.ch	5%
news.tirol.com	4%
www.bielertagblatt.ch	4%
www.fr-aktuell.de	4%
www.nachrichten.at	4%
www.netecho.info	4%
Gesamt	100%

Abb. 13: 50 Prozent Cut Off Paperball



3.5.2 Kommentar zu den Ergebnissen

Die Häufigkeit der Treffer beim *Hamburger Abendblatt* mag seinen Grund in der stärkeren lokalen Berichterstattung zum Thema Hamburger Innensenator Schill haben. Interessant ist hingegen die Tatsache, dass die Suchdienste bei benachbarten Angeboten wildern: AltaVista liefert die (zur Unternehmensgruppe gehörenden) Yahoo-News am häufigsten. Außerdem fällt auf, dass bei Agenturmeldungen oft kleine Zeitungen gefunden werden, da diese Agenturmeldungen einfach automatisch und unredigiert übernehmen.

3.6 Ergebnisdiskussion und aktueller Stand der Nachrichtensuche

Die Konzentrationstendenzen der Suchmaschinen betreffen auch die Nachrichtensucher. Es gibt nur drei eigene, ernst zu nehmende Nachrichtensucher, die mehrere, voneinander unabhängige Quellen durchsuchen. Des Weiteren finden sich am Markt zahlreiche Angebote, die lediglich Newsfeeds verschiedener Agenturen durchreichen.

Wer sich nur aus einer Internet-Nachrichtenquelle versorgt, verpasst zwangsläufig eine große Zahl von Nachrichten und Meinungen. Auch die Nutzung mehrerer Internet-Quellen garantiert jedoch keine vollständige Wiedergabe der Nachrichtenlage. Andererseits bieten erst die Internet-Nachrichtensucher überhaupt einen breiten Zugriff auf die unterschiedlichen Nachrichtenquellen. Ein Zugriff auf derart viele Nachrichten unterschiedlichster

Art nur mit Print-Erzeugnissen könnte ohne erheblichen Ressourcenaufwand nicht stattfinden.

Generell dominieren im Nachrichtenangebot des Internets aktuelle Meldungen. Hintergrundberichte, Analysen und Kommentare erscheinen kaum. Grund hierfür ist einerseits die Tatsache, dass entsprechend gründlich recherchierte Berichte selten den Weg ins kostenlose Internet finden. Andererseits sorgt der sich ständig ändernde Nachrichtenfluss dafür, dass, sollten entsprechende Veröffentlichungen in den Nachrichtensuchern auftauchen, diese schnell von weiteren Meldungen verdrängt werden. Da die Zeitungen und Zeitschriften nur einen Bruchteil ihrer Artikel, die sie in ihren Print-Ausgaben publizieren, online zur Verfügung stellen, kann auch aus diesem Grunde kein umfassender Presseüberblick gegeben sein. Somit sind aktuelle Meldungen oftmals alles, was im Netz kostenlos durch Suchmaschinen auffindbar ist. Die Nachrichtenbeschaffung via Nachrichtensuchmaschinen entspricht der auch in anderen Medien anzutreffenden gegenwärtigen Tendenz, laufend aktuelle Informationen möglichst schnell zu erhalten – ohne weitere Einsortierung in einen Wissenskontext.

Nachrichtensucher haben einen vergleichsweise hohen technischen Stand erreicht: Es ist möglich, definierte Datenströme aus strukturierten Dokumenten zeitnah zu bündeln, zu aggregieren und durchsuchbar zu machen. Die technischen Möglichkeiten dürfen aber nicht darüber hinwegtäuschen, dass die Qualität der Suchergebnisse immer inhaltlich, wirtschaftlich und technisch limitiert ist:

1. Technisch besser erfassbare Webangebote können überrepräsentiert sein.
2. Wirtschaftliche Gründe können Anlass zu vermindertem Angebotsspektrum und verminderter Quellenauswahl sein. Die Pflege der Nachrichtensuche ist kostenträchtig.
3. Inhaltliche Vorlieben der Redaktion bei der Quellenauswahl können die Ergebnisse beeinflussen.

Da Nachrichtensuchmaschinen keinesfalls kostendeckend arbeiten, sondern von den Betreibern aus anderen Erlösquellen querfinanziert werden, ist das Angebot wirtschaftlichen Zwängen insofern unterworfen, als dass so preiswert wie nur möglich gearbeitet werden muss, und der Endnutzer immer auf den guten Willen des Anbieters angewiesen ist.

Auf Anbieterseite sind Nachrichtensuchmaschinen weniger als Einkommensquelle als vielmehr als Mittel zur Kundenbindung zu sehen: Nutzer, die zur Nachrichtensuche ein Angebot aufsuchen, kommen mit großer Wahrscheinlichkeit auch zur Websuche wieder und erhöhen somit die Marktanteile und damit auch den Marktwert einer Suchmaschine. Das Fehlen einer Nachrichtensuche könnte zudem eine Abwanderung von Nutzern bewirken.

4. Ausblick

Die Nachrichtensuche ist ein teures Geschäft, für das gegenwärtig kein Verdienstmotiv existiert. Die übliche Praxis, bezahlte Treffer mitauszuliefern, würde bei Nachrichten zu bizarren Ergebnissen führen – kommerzielle Anzeigen müssten neben Katastrophennachrichten platziert werden.

Jeder Anbieter einer Nachrichtensuche hat bei seiner Umsetzung freie Hand und gestaltet das Angebot nach seinem Geschmack – entsprechend sind Verzerrungen in der Nachrichtenauswahl denkbar, wenn auch derzeit nicht nachweisbar. Die Nachrichtensuche wird im Vergleich zur normalen Websuche deutlich weniger genutzt, obwohl die aktuellen Nachrichten teilweise in die normalen Suchergebnisse eingeblendet werden.

Zu beobachten ist, dass die amerikanischen Dienste ihre Quellen nicht transparent machen, und so nicht nachvollziehbar ist, welche Publikationen gar nicht berücksichtigt werden oder nur nicht so gut „gerankt“ werden wie andere. Gerade vor dem Hintergrund des oligopolartig aufgeteilten Suchmaschinenmarktes²⁸ sind derartige Tendenzen bedenklich im Hinblick auf Meinungsfreiheit und -vielfalt zu bewerten.

Festzustellen ist auch, dass die Ticker-Meldungen der großen Nachrichtenagenturen das Feld weitgehend bestimmen, Kolumnen, Feature-Artikel oder ausführlichere Berichte verweilen nur kurz an der Spitze der Trefferlisten und werden schnell von neueren Meldungen verdrängt. Analog zur Entwicklung in anderen Medien ist die Verbreitung der aktuellsten Meldung vor dem Hintergrundbericht deutlich vorrangig. Meinung entsteht weniger aus reflektiertem Wissen als aus der Ad-hoc-Meldung – allzu oft ohne Hintergrund und Zusammenhang.

Die Nachrichtensuche könnte durch Einsatz weiterer Technik- und Personalressourcen optimiert werden – dies ist wegen mangelnder Refinanzierungsmöglichkeiten jedoch unwahrscheinlich. Um die aktuelle Nachrichtenlage auch in der normalen Websuche klarer zu kommunizieren, ist eine nahtlose Einbindung von Nachrichten in die Suchergebnisse denkbar. Derzeit existiert bei Google und Yahoo nur eine Anzeige von Nachrichten-Links über den normalen Suchergebnissen.

Die explizite Nachrichtensuche wird im Verhältnis zur normalen Suche nur von einem kleinen Anteil von Suchmaschinennutzern verwendet. Je nahtloser die Ergebnisse in der normalen Trefferliste präsentiert würden, desto höher wäre die Durchdringung der Alltagssuche mit dem aktuellen Nachrichtenstand.

28 So dominieren nach Nielsen-Netrating die drei großen Suchmaschinen (Google, Yahoo, MSN) in den USA den Suchmaschinenmarkt deutlich (vgl. Sullivan 2004). Allein Google wird von knapp 40 Prozent der Nutzer für Suchanfragen verwendet. Yahoo und MSN kommen immerhin bei jeweils rund 30 Prozent der Nutzer zum Einsatz.

Das Problem der Zuordnung einzelner Nachrichten in einen Gesamtkontext und damit der Schritt von der schlichten Information hin zum gesellschaftspolitischen Wissen, das aus verknüpften aktuellen und zeitlosen Informationen besteht, ist von Internet-Nachrichtensuchern ohne redaktionelle Zuarbeit nicht zu leisten – die Meinungsbildung über den Informationskanal Internet ist auch weiterhin auf die Medienkompetenz des Einzelnen angewiesen. Ein Grund mehr, um auch hier Transparenz zu fordern.

URLs der Nachrichtensuchmaschinen

Alltheweb:	http://www.alltheweb.com/?cat=news
AltaVista:	http://de.altavista.com/news/
AOL:	http://suche.aol.de/suche/index.jsp http://www.aol.de/index.jsp?sg=News
Fireball:	http://www.fireball.de
Google News:	http://news.google.de/
Lycos:	http://www.lycos.de/startseite/news/
Metager:	http://www.metager.de
MSN Newsbot:	http://uk.newsbot.msn.com/
Paperazzi:	http://www.paperazzi.de
Paperball:	http://paperball.fireball.de/
T-Online:	http://brisbane.t-online.de/fast-cgi/tsc?PortalLanguage=de&UserLanguage=de&context=profi-search&device=html&mandant=toi&q=&search_type=profi&tpc=service&active_tab=news
Yahoo:	http://de.news.yahoo.com/
WEB.DE:	http://portale.web.de/schlagzeilen/

Literatur

- Bekavac, Bernard (2002): Methoden und Verfahren von Suchdiensten im WWW/Internet.
http://www.inf-wiss.uni-konstanz.de/suche/tutorial/such_tutorial_advanced.html
(10.6.2004).
- Bentele, Günter (1999): Parasitismus oder Symbiose? Das Intereffikationsmodell in der Diskussion. In: Lothar Rolke/Volker Wolff (Hrsg.): Wie die Medien die Wirklichkeit steuern und dabei selber gesteuert werden. Opladen/Wiesbaden. S. 177-194.
- Bentele, Günter (2003): Kommunikatorforschung: Public Relations. In: Ders. (Hrsg.): Öffentliche Kommunikation. Wiesbaden, S. 54-78.
- Bertelsmann-Stiftung: Die Kluft zwischen Wissenschaft und Öffentlichkeit überwinden. Without Author.

- http://www.bertelsmann-stiftung.de/medien/pdf/forum_4_03_Wissenschaftsjournalismus.pdf (21.10.2004).
- BGH (2003): Bundesgerichtshof, Urteil vom 17. Juli 2003, Aktenzeichen: I ZR 259/00.
- Blöbaum, Bernd (1994): Journalismus als soziales System. Geschichte, Ausdifferenzierung und Verselbständigung. Opladen.
- Bondolfi, Alberto (1999): Die Pflicht zur Wahrheitsaussage. In: Holderegger, Adrian: Kommunikations- und Medienethik. Freiburg, S. 277-298.
- Borchers, Detlef (2003): Worte Waschen-Wer bestimmt, was Google findet? In: Neue Zürcher Zeitung, 11. April 2003.
- Brin, Sergey/Page, Lawrence (1998): The Anatomy of a Large-Scale Hypertextual Web Search Engine. Stanford.
<http://www-db.stanford.edu/~backrub/google.html> (16.3.2004).
- Bruckenberg, Johannes/ Mayr, Wolfgang (2004): Journalismus in Nachrichtenagenturen. In: Heinz Pürer, Meinrad Rahofer, Claus Reitan (Hrsg.): Praktischer Journalismus. Presse, Radio, Fernsehen, Online. Konstanz, S. 161-168.
- Calishain, Tara (2002): Monitoring the News Online. In: Online Vol. 26, No. 4, pp.24-28.
- Chu, Heting (2003): Information Representation and Retrieval in the Digital Age. Medford, NJ: Information Today.
- Diekhof, Rolf (2003): Der Gummibärchenmann. In: Zukunft der Werbung – Werbung der Zukunft. Ein gemeinsames Supplement zum Deutschen Werbekongress 2003 von w&v werben und verkaufen, Süddeutsche Zeitung, media und marketing und Der Kontakter, S. 6-7.
- Eilders, Christiane (1997): Nachrichtenfaktoren und Rezeption. Eine empirische Analyse zur Auswahl und Verarbeitung politischer Information. Opladen.
- Finkelstein, Seth: Google Bayesian Spam Filtering Problem?
<http://sethf.com/anticensorware/google/bayesian-spam.php> (26. November 2003).
- Franck, Georg (1998): Ökonomie der Aufmerksamkeit. Ein Entwurf. München: Hanser.
- Fröhlich, Romy (1992): Qualitativer Einfluss von Pressearbeit auf die Berichterstattung: Die geheime Verführung der Presse? In: Publizistik, Vol. 37, Nr. 1, S. 37-49.
- Galtung, Johan/Ruge, Mari Holmboe (1965): Structure of Foreign News. The Presentation of the Congo, Cuba and Cyprus Crises in Four Norwegian Newspapers, in: Journal of Peace Research 2, S. 64-91.
- Goldhaber, Michael H. (27.11.1997): Die Aufmerksamkeitsökonomie und das Netz – Teil I und Teil II. Übersetzt von Florian Rötzer.
<http://www.ct.heise.de/tp/deutsch/special/eco/6195/1.html> (Teil I, 16.1.2004)
<http://www.ct.heise.de/tp/deutsch/special/eco/6200/1.html> (Teil II, 16.1.2004).
- Hagen, Lutz (1995): Informationsqualität von Nachrichten. Opladen.
- Jungbluth, Rüdiger (2003): Wer googelt der findet. In: Max, März 2003, S. 128-133.
- Karzauninkat, Stefan (2003): Die Suchmaschinenlandschaft 2003: Wirtschaftliche und technische Entwicklung. In: Machill, Marcel/Welp, Carsten (Hrsg.): Wegweiser im Netz. Qualität und Nutzung von Suchmaschinen. Gütersloh, S. 509-538.
- Kepplinger, Mathias (2001): Der Ereignisbegriff in der Publizistikwissenschaft. In: Publizistik Heft 2, S. 117-139.

- Klußmann, Niels (2001): Lexikon der Kommunikations- und Informationstechnik: Telekommunikation, Internet, Mobilfunk, Multimedia, Computer, E-Business. Heidelberg.
- Krumenaker, Larry (2003): Tribunes and Tribulation: The Top 100 Newspaper Archives (or Lack Thereof). In: Searcher, Vol. 11, Nr. 7, S. 28-35.
- Lewandowski, Dirk (2004): Technologie-Trends im Bereich der WWW-Suchmaschinen. In: Information Professional 2011: Proceedings der 26. Online-Tagung der Deutschen Gesellschaft für Informationswissenschaft. Frankfurt am Main, 2004 [i. Dr.].
- Machill, Marcel/Neuberger, Christoph/Schindler, Friedemann (2002): Transparenz im Netz. Funktionen und Defizite von Internet-Suchmaschinen. Gütersloh
- Machill, Marcel/Neuberger, Christoph/Schweiger, Wolfgang/Wirth, Werner (2003): Wegweiser im Netz. Qualität und Nutzung von Suchmaschinen. In: Machill, Marcel/Welp, Carsten (Hrsg.): Wegweiser im Netz. Qualität und Nutzung von Suchmaschinen. Gütersloh 2003, S. 13-490.
- McHugh, Josh: Google vs. Evil. In: Wired Magazine, 1/2003.
- Medientipp.ch (5.10.1999): Das journalistische Gewissen und Qualität im Journalismus. Without Author.
<http://www.medientipp.ch/index.php?na=4,4&meid=32090&pv=y> (21.10.2004).
- Neuberger, Christoph (2003): Online-Journalismus: Veränderungen – Glaubwürdigkeit – Technisierung. In: Media-Perspektiven 3/2003, S. 131-138.
- Noelle-Neumann, Elisabeth/Mathes, Rainer (1987): The „Event as Event“ and the „Event as News“: The Significance of „Consonance“ for Media Effects Research. In: European Journal of Communication 2, S. 391-414.
- Notess, Greg R. (2002): The Blog Realm: News Sources, Searching with Daypop, and Content Management. In: Online Vol. 26, Nr. 5, S. 70-2.
- Notess, Greg R. (2003): Search Engine Statistics: Freshness Showdown.
<http://www.searchengineshowdown.com/stats/freshness.shtml> (10.6.2004).
- Page, Lawrence/Brin, Sergey/Motwani, Rajcew/Winograd, Terry (1999): The Page-Rank Citation Ranking: Bringing Order to the Web.
<http://dbpubs.Stanford.edu/pub/showdoc.fulltext?lang=en&doc=1999-66&format=pdf&compression=> (24.3.2004).
- Requate, Jörg (1995): Journalismus als Beruf. Göttingen: Vandenhoeck und Ruprecht.
- Rösch, Hermann (2001a): Vom Suchwerkzeug zum Portal. In: Password, Vol. 16, Nr. 3, S. 18-25.
- Rösch, Hermann (2001b): Portalfunktionen und typologische Varianten. In: Password Vol. 16, Nr. 4, S. 26-35.
- Rötzer, Florian: Die Welt ist keineswegs alles, was Google auflistet.
<http://www.heise.de/tp/deutsch/inhalt/te/13486/1.html> (25.10. 2002).
- Ruß-Mohl, Stephan (1994): Der I-Faktor: Qualitätssicherung im amerikanischen Journalismus – Modell für Europa? Zürich .
- Shoemaker, Pamela J. (1991): Communication Concepts 3: Gatekeeping. Newsbury Park, Calif. 1991.
- Schiff, Frederick (2003): Business Models of News Web Sites: A Survey of empirical trends and expert opinion. In: First Monday Vol. 8, Nr. 6.
http://firstmonday.org/issues/issue8_6/schiff/index.html (17.11.2003).

- Schönhausen, Philomen (1998): Unparteilichkeit im Journalismus. Tradition einer Qualitätsnorm. Tübingen.
- Schröter, Detlef (1995): Qualität und Journalismus. Theoretische und praktische Grundlagen journalistischen Handelns. München.
- Schulz, Winfried (1999): Nachricht. In: Noelle-Neumann, Elisabeth/Schulz, Winfried/Wilke, Jürgen: Fischer Lexikon. Publizistik der Massenkommunikation. Frankfurt/Main: Fischer, S. 307-337.
- Schwiesau, Dietz/Ohler, Josef (2003): Die Nachricht in Presse, Radio, Fernsehen, Nachrichtenagentur und Internet. Ein Handbuch für Ausbildung und Praxis. München.
- Spencer, Stephan M. (2002): The Search Is On Catalog Age, 1. Juni.
<http://catalogagemag.com/ar/marketing-search/> (16.3.2004).
- Sullivan, Danny (2003): News Search Engines.
<http://www.sewatch.com/links/article.php/2156261> (26.11.2003).
- Sullivan, Danny (2004): Nielsen NetRatings Search Engine Ratings.
<http://searchenginewatch.com/reports/article.php/2156451> (14.6.2004).
- Van Eimeren, Birgit/Gerhard, Heinz/Frees, Beate (2003): ARD/ZDF-Online-Studie 2003: Internetverbreitung in Deutschland: Unerwartet hoher Zuwachs. In: Media Perspektiven, Nr. 8, S. 338-358.
- Weaver, David H. (Ed. 1998): The Global Journalist. Cresskill: Hampton Press.
- Weischenberg, Siegfried (1985): Die Unberechenbarkeit des Gatekeepers. Zur Zukunft professioneller Informationsvermittlung im Prozess technisch-ökonomischen Wandels. In: Rundfunk und Fernsehen, Nr. 2, S. 187-201.
- Weischenberg, Siegfried (1995): Journalistik. Medientechnik, Medienfunktionen, Medienakteure. Bd. 2. Opladen: Westdeutscher Verlag.
- Weischenberg, Siegfried (2001): Nachrichten-Journalismus. Anleitungen und Qualitäts-Standards für die Medienpraxis. Wiesbaden: Westdeutscher Verlag.
- Weischenberg, Siegfried (2002): Apocalypse how? Journalismus und Patriotismus. (Statement at: 1. Petersberger Kongress „Medien – Instrumente des Terrorismus?“, 29. August 2002).
- Welp, Carsten (2003): Ein Code of Conduct für Suchmaschinen. In: Machill, Marcel/Welp, Carsten (Hrsg.): Wegweiser im Netz. Qualität und Nutzung von Suchmaschinen. Gütersloh 2003, S. 491-508.
- White, David Manning (1964): The „Gatekeeper“. A Case Study in the Selection of News. In: Lewis A. Dexter/David M. White (Hrsg.): People, Society and Mass Communications, London, S. 160-172.
- Wiggins, Richard W. (2001): The Effects of September 11 on the Leading Search Engine. First Monday, Vol. 7, Nr. 10.
http://www.firstmonday.dk/issues/issue6_10/wiggins/index.html (17.11.2003).
- Wimmeroth, Ulrich/Brochhagen, Thomas (2003): Dirty Tricks – Google. Düsseldorf: Data Becker.
- Zerdick, Axel et al. (2001): Die Internet-Ökonomie. Strategien für die digitale Wirtschaft (European Communication Council Report). Berlin, Heidelberg, New York.
- Zurstiege, Guido/Siegfried J. Schmidt (2003): Werbekommunikation. In: Günter Bentele (Hrsg.): Öffentliche Kommunikation. Wiesbaden, S. 492-503.

Anhang

Die im Rahmen dieser Arbeit betrachteten Nachrichtensuchmaschinen und News-Portale sind nicht für die Recherchearbeit von Journalisten gedacht, sondern bieten eher eine Informationsfunktion für private Internet-Nutzer. Es gibt allerdings auch Nachrichtenangebote, die sich speziell an Journalisten richten und als Informationsquellen und Arbeitshilfen gedacht sind. Die folgende Zusammenstellung soll einen kurzen, nicht erschöpfenden Überblick bieten. Auch die Netzangebote der wichtigsten deutschsprachigen Nachrichtenagenturen werden mit ihren Funktionen vorgestellt.

Speziell für Journalisten konzipierte Nachrichtenangebote

Journalistenlinks.de

<http://www.journalistenlinks.de>

Ein Portal mit Links zu Pressemitteilungen, Medienseiten und einem Katalog mit Links, die die Arbeit von Journalisten erleichtern sollen. Die Suche führt zu Recherchequellen und Hintergrundinformationen, nicht zu aktuellen Meldungen. Journalistenlinks versteht sich als „Webkatalog von und für Journalisten“.

Newsroom

<http://www.newsroom.de>

Ein Portal mit aktuellen News der Agenturen dpa, AFP und ddp. Die Meldungen für Nichtmitglieder werden erst nach zwölf Stunden veröffentlicht. Das Archiv präsentiert die Ticker-Meldungen in chronologischer Reihenfolge.

textinternet

<http://www.textintern.de>

textinternet bringt Meldungen aus der Medienbranche, normale News der Nachrichtenagenturen sind nicht zu finden.

Journalisten & Journalismus

<http://www.journalismus.com>

„Der Journalisten-Treff im Web“ ist ein umfassendes Portal mit schwerpunktmäßig für Journalisten interessanten Themen. Die Suchfunktion durchsucht die Site, nicht die Nachrichtenmeldungen. Der OTS-Ticker sowie ein Branchen-Ticker sind eingebunden. Dazu gibt es Foren, und einen Journalismuskatalog mit etwa 3.500 Einträgen.

Jourweb

<http://www.jourweb.com>

Eine Kollektion unterschiedlicher, auch internationaler News-Ticker und Links zu Nachrichtenquellen, Pressestellen und Medienarchiven, nach Themengebieten gruppiert.

Presse-Suche

<http://www.presse-suche.de>

Eine Suchmaschine, die sich im Durchsuchen von Medienseiten versucht. Der Robot unterscheidet weder zwischen Nachrichten und anderen Seiten – listet beispielsweise Immobilienanzeigen und Übersichtsseiten – , noch werden die zentralen Inhalte einer Seite extrahiert.

Pressetreff

<http://www.pressetreff.de>

Der Betreiber Deutsche Journalisten Dienste stellt eingetragenen Mitgliedern Pressetexte zu nichtaktuellen, allgemeinen Themen zur Verfügung.

newsclub.de

<http://www.newsclub.de>

Eine sehr schnell arbeitende, nach Kategorien sortierte News-Meta-Suchmaschine mit Angabe der Quelle. Die Treffer werden nur als Headlines ohne Abstract präsentiert und lassen sich nach Uhrzeit, Quelle und alphabetisch sortieren.

Recherchetipps.de

<http://www.recherchetipps.de>

Ein reiner Katalog für journalistische Rechercheaufgaben; es sind keine Agentur-Ticker eingebunden.

Onlinejournalismus

<http://www.onlinejournalismus.de>

Ein E-Zine²⁹ zum Thema Online-Journalismus mit vierzehntägig erscheinenden Artikeln und einem gut besuchten Forum. Keine Nachrichten-Ticker.

News-Ticker.org

<http://www.news-ticker.org>

Eher zweifelhaftes Angebot, macht mit Spam Mails auf sich aufmerksam und bringt nur Pressemeldungen. Betreiber ist ein Fotolabor.

Zusammenfassung

Die Journalistenportale bieten – neben branchenspezifischen Informationen – durchgereichte Ticker-Meldungen der üblichen Agenturen sowie einige wenige eigene Artikel. Die Recherchemöglichkeiten und die Suche reichen nicht an die Möglichkeiten spezieller Nachrichtensuchmaschinen oder gar profes-

29 Electronic-Magazine – Bezeichnung für eine reine Online-Publikation.

sioneller Pressedatenbanken heran. Für die Online-Recherche von Journalisten stellen die vorgestellten Angebote aber eine Arbeitserleichterung dar. Die Funktion der Seiten als Orientierungshilfe und Wegweiser zu relevanten Seiten ermöglicht zielgerichtetes Arbeiten ohne großen Zeitverlust bei der Suche nach verwertbaren Informationen.

Nachrichtenagenturen

Reuters

<http://www.reuters.de/>

Die Nachrichtenagentur betreibt ein eigenes Portal, in dem die Meldungen aufbereitet werden. Eine Suchfunktion ist nicht vorgesehen. Das Archiv listet chronologisch etwa 30 Meldungen pro Seite, sortiert nach den Rubriken „Top-Nachrichten, Politik, Wirtschaft, Unternehmen“, und „Panorama“.

AFP

<http://www.afp.com/deutsch/home/>

Die Website informiert in erster Linie über Produkte und Dienstleistungen der internationalen Nachrichtenagentur. Ticker-Meldungen werden nur beiläufig eingeblendet. Es gibt weder ein Archiv noch eine Suchmöglichkeit.

dpa

<http://www.dpa.de>

Die dpa-Website liefert keinerlei Nachrichten, es handelt sich um eine reine Corporate Website, die über Produkte und Dienstleistungen informiert.

Presseportal

<http://www.presseportal.de>

Portal der dpa-Pressemeldungen. Gegen Gebühr kann jeder über den Presseverteiler seine Pressemeldungen verschicken. Das Presseportal liefert ausschließlich Pressemeldungen von Firmen und Organisationen, keine normalen Nachrichten.